

ABSTRACT

Title of Document: ESTIMATING UNKNOWN KNOTS IN PIECEWISE LINEAR-LINEAR LATENT GROWTH MIXTURE MODELS

Nidhi Kohli, Doctor of Philosophy, 2011

Directed By: Dr. Gregory R. Hancock, and Dr. Jeffrey R. Harring, Department of Measurement, Statistics & Evaluation

A piecewise linear-linear latent growth mixture model (LGMM) combines features of a piecewise linear-linear latent growth curve (LGC) model with the ideas of latent class methods all within a structural equation modeling (SEM) context. A piecewise linear-linear LGMM is an appropriate framework for analyzing longitudinal data that come from a mixture of two or more subpopulations (i.e., latent classes) where each latent class incorporates a separate growth trajectory corresponding to multiple growth phases from which repeated measurements arise. The benefit of the model is that it allows the specification of each growth phase to conform to a particular form of overall change process within each latent class thereby making these models flexible and useful for substantive researchers.

There are two main objectives of this current study. The first objective is to demonstrate how the parameters of a piecewise linear-linear LGMM, including the unknown knot, can be estimated using standard SEM software. A series of Monte Carlo simulations empirically investigated the ability of piecewise linear-linear LGMMs to

recover true (known) growth parameters of distinct populations. Specifically, the current research compared the performance of the piecewise linear-linear LGMM under different manipulated conditions of 1) sample size, 2) class mixing proportions, 3) class separation of location of knot, 4) the mean of the slope growth factor of the second phase, 5) the variance of the slope growth factor of the second phase, and 6) residual variance of the observed variables.

The second objective is to address the issue of model mis-specification. It is important to analyze this issue because applied researchers have to make model selection decisions. Therefore, the current research examined the possibility of extracting spurious latent classes. To achieve this objective 1-, 2-, and 3-class piecewise linear-linear LGMMs were fit to data sets generated under different manipulated conditions using a 2-class piecewise linear-linear LGMM as a population model. The number of times the correct model (i.e., 2-class piecewise linear-linear LGMM) was preferred over incorrect models (i.e., 1- and 3-class piecewise linear-linear LGMMs) using the Bayesian Information Criterion (BIC) was examined.

Results suggested that the recovery of model parameters, specifically, the variances of growth factors were generally poor. In addition, none of the manipulated conditions were systematically related to the outcome measures, parameter bias and variability index of parameter bias. Furthermore, among all the manipulated conditions, the residual variance of observed variable had the strongest statistically significant effect on both the model convergence rate and the model selection rate. Other manipulated conditions that had an impact on the model convergence rate and/or the model selection rate were the growth factor mean of slope of the second phase, the growth factor variance

of slope of the second phase, and the class mixing proportion. The manipulated conditions whose levels had no influence on either the model convergence rate or the model selection rate were sample size and the class separation of location of knot.

ESTIMATING UNKNOWN KNOTS IN PIECEWISE LINEAR-LINEAR
LATENT GROWTH MIXTURE MODELS

By

Nidhi Kohli

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2011

Advisory Committee:

Professor Gregory R. Hancock, Co-Chair
Professor Jeffrey R. Harring, Co-Chair
Professor Hong Jiao
Professor George Macready
Professor Paul Hanges

© Copyright by
Nidhi Kohli
2011

DEDICATION

To Mumma and Papa, for your unconditional love and support. Had it not been for your constant encouragement and unwavering faith in me, I would never have been able to make it this far in my education. To my husband Vaibhav, for your love, patience and immense support, and for going through this five year journey with me. To Nitin, my big brother, for your love and encouragement, and without whom I would probably have never come to the U.S. to pursue my advanced education. Thank you all for helping me to grow, I love each of you.

ACKNOWLEDGEMENTS

This work would never have been possible without the inspiration and support from a lot of people. I know for sure however sincere an effort I might make to acknowledge the contributions of everybody who helped me through this, it will always fall short of what actually they mean to me.

I would first like to thank my advisor, Dr. Gregory R. Hancock, for his constant support and encouragement during my years at University of Maryland. You always believed in me and in my ability to succeed in the PhD program. You are an outstanding academic advisor and dissertation co-chair and I have learned so much under your tutelage of which I can be proud.

I would also like to thank Dr. Jeffrey R. Haring, my dissertation co-chair, for his continued guidance and help and giving me the opportunity to work on some very interesting projects. Your critical insight and experience has greatly added to my skills, knowledge, and confidence needed to become an independent researcher.

My sincere thanks also go to Dr. George Macready, Dr. Hong Jiao, and Dr. Paul Hanges for agreeing to serve on my committee and providing constructive comments. Lastly, I want to thank Vaibhav Dua for your support and being my savior whenever I got stuck in the analytical programming related to my dissertation; Vinod and Meenakshi Kohli, and Nitin Kohli for constantly motivating me to continue to pursue my PhD dream at times when I couldn't motivate myself enough.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION	1
CHAPTER 2: REVIEW OF LITERATURE	6
2.1 Latent Growth Curve Models.....	6
2.1.1 ML estimation of latent growth curve models.....	11
2.2 Piecewise Latent Growth Curve Models.....	12
2.2.1 ML estimation of piecewise linear-linear latent growth curve models with unknown knot	20
2.3 Finite Mixture Models.....	21
2.3.1 Mixtures of univariate distributions	22
2.3.2 Mixtures of multivariate distributions	24
2.3.3 Latent growth mixture models.....	24
2.3.4 ML estimation of latent growth mixture models via the EM Algorithm.....	27
2.4 Research Goals	30
CHAPTER 3: METHODOLOGY	32
3.1 Research Design.....	32
3.2 Data Generation Model	34
3.3 Simulation Design.....	36
3.2.1 Fixed conditions	37
3.2.1 Manipulated conditions	38
3.4 Pilot Analysis	44
3.4.1 Selection of Levels of Sample Size Condition	44
3.4.2 Asymptotic Behavior of Model Parameters for a 2-Class Piecewise Linear- Linear LGMM	45
3.5 Model Estimation	46
3.6 Outcome Measures	47
CHAPTER 4 – RESULTS	50
4.1 Model Convergence Rate	50
4.3. Model Mis-specification	59
4.3.1. Model selection rate.....	60

4.3.2. Parameter bias and variability index of parameter bias for filtered replications	63
4.4. Summary of Main Findings	64
CHAPTER 5 – DISCUSSION	66
5.1 Summary of Results	67
5.1.1 Model convergence rate.....	67
5.1.2 Parameter bias.....	68
5.1.3 Model selection rate.....	69
5.2 Limitations of Study.....	70
5.3 Recommendations for Applied Researchers	71
5.4 Methodological Extensions	73
APPENDIX – A.....	75
APPENDIX – B	77
APPENDIX – C	79
APPENDIX – D.....	87
REFERENCES	93

LIST OF TABLES

Table 1:	Population parameter values for fixed conditions.....	38
Table 2:	Manipulated conditions.....	39
Table 3:	Population parameter values for fixed and manipulated.....	43
Table 4:	Proportion of properly converged replications for two-class piecewise linear-linear LGMMs.....	51
Table 5:	Percentage of conditions with 10% or more bias in either direction.....	58

LIST OF FIGURES

Figure 1:	A conventional first-order linear LGC model with linear trajectory and mean factor.....	8
Figure 2:	A spaghetti plot depicting generic piecewise linear-linear change process.....	13
Figure 3:	A conventional piecewise linear-linear LGC model where the location of knot is <i>known</i>	16
Figure 4:	A conventional piecewise linear-linear LGC model where the location of knot is <i>unknown</i>	18
Figure 5:	A two-class univariate normal mixture model.....	23
Figure 6:	A multivariate normal mixture model.....	24
Figure 7:	A conventional two class linear LGMM.....	27
Figure 8:	Classification Tree: The effect of manipulated conditions on the percentage of properly converged replications.....	54
Figure 9:	Classification Tree: The effect of manipulated conditions on the selection rate.....	61

CHAPTER 1: INTRODUCTION

A common challenge for researchers and practitioners across different research domains is to understand how individuals change and develop on certain variables over time. For instance, when new skills are acquired, or when attitudes and interests develop, people change. Measuring change over time requires a longitudinal perspective, where repeated measurements are gathered for a sample of individual subjects. Several statistical methods have been developed in the past three decades to analyze longitudinal data of this type, including mixed effects models (Laird & Ware, 1982), multilevel models (Goldstein, 2003), and latent growth curve models (Meredith & Tisak, 1990).

Mixed effects models for longitudinal data, including both linear and nonlinear mixed effects models, can be viewed as a generalization of the conventional multiple regression models. At its core, linear mixed effects models are used when individual growth trajectories show straight-line patterns of change. The model characterizes the change process by a function common to all subjects, but whose parameterization is allowed to vary among individual subjects, thereby allowing for between-subject variability. When the response pattern is, however, distinctly nonlinear one option is to extend the linear mixed effects models by incorporating higher-order polynomial terms to account for the curvilinear time-response pattern. Another option is to characterize the nonlinear change process with an intrinsically nonlinear function (i.e., the response function has at least one parameter that enters nonlinearly).

The linear mixed effects model is a subject-specific model, and as such, the main focus of analyses using this model is on change at the individual-subject level, rather than

average change aggregated across individuals. Longitudinal models stemming from a factor analytic tradition, the latent growth curve (LGC) model is based on the idea that the process of change over time in repeated measures data is described by an underlying latent process. The LGC model is defined for each individual subject; however, the main focus of the analysis is on the change at the population level, that is, average growth trajectory in the population rather than change at the individual-subject level (Cudeck & Haring, 2007). Moreover, because LGC models reside in the SEM family, the LGC approach has the advantage of providing measures of model fit statistics that enable researchers to conduct hypothesis testing (e.g., Chou, Bentler, & Pentz, 1998). In contrast, in the mixed effects modeling approach, there is no single inferential test of the overall goodness-of-fit of a specific hypothesized mixed effects model (Curran, 2003). On the other hand, an advantage of using mixed effects modeling approach is that it is very flexible and efficient analytic framework (e.g., Chou et al., 1998; Curran, 2003). The mixed effects modeling of nested or hierarchical data structure involves the simultaneous disaggregation of the level-1 and level-2 covariance structures, where the disaggregated effects can be estimated by including predictors in either level-1 or level-2 parts of the model. In contrast, the estimation of LGC models is based on a single aggregate covariance matrix that allows for covariance structure only at a single level of analysis; the covariance structure within any other level of nesting is assumed to be null (Curran, 2003). There is a great deal of overlap between these two modeling frameworks, though. As Curran (2003) stated, “Indeed, the boundaries between these two modeling strategies are becoming increasingly porous as is evidenced in that fully random regressions can be

estimated in the SEM and latent variable measurement models can be estimated in the multilevel model framework” (p. 565).

For the purpose of this study, the derivations, implementation, and subsequent discussions, as well as the extensions pertaining to the statistical method for analyzing longitudinal data will be done in a SEM framework (i.e., the LGC model and its extensions). One reason that an SEM framework may be preferable is that maximum likelihood (ML) estimation used in SEM programs which focuses on minimization of a discrepancy function between observed and model-implied covariance matrices, often quickly converges to a solution. Another compelling reason is that SEM programs, for many applied and methodological researchers, are the analytic tools of choice for most purposes. Thus, staying within a familiar software environment is not only convenient, but it also avoids the inefficiencies involved in setting up a specific model correctly using an unfamiliar syntax language. The estimation of LGC models within a SEM framework is discussed in detail in Chapter 2.

The LGC model is a popular and relatively simple technique for modeling change at both the individual and population level. The model consists of continuous latent growth factors, which correspond to the average growth trajectory characteristics in the population. The LGC model also allows the variability of the repeated measures to be partitioned into within-subject and between-subject components. In LGC models, it is typically assumed that the functional form describing the overall change process in the repeated measures data is smooth and continuous. However, assuming that a single uninterrupted functional form underlying the overall change process may be unrealistic for applications where data are comprised of different growth phases.

The piecewise latent growth curve model, an extension of the LGC model, allows the specification of each growth phase to conform to a particular functional form of the overall change process (Chou, Yang, Pentz, & Hser, 2004; Cudeck & Haring, 2010). An interesting feature of a piecewise LGC model is the location of the knot(s) (or changepoint). The location of the knot in a two-phase change process is the value of the predictor at which the function shifts from one phase to the other (Cudeck, 1996; Cudeck & Klebe, 2002), and its location can be known a priori or estimated. The estimation of the knot in piecewise LGC models is also discussed in detail in Chapter 2.

An assumption inherent to the LGC model is that all individuals are drawn from the same population, and thus share the same functional form of growth. This assumption, which applies to the piecewise LGC model as well, is not practical in situations where the data come from a mixture of two or more unobserved subpopulations (i.e., latent classes). The analysis of this type of mixture data requires the expansion of the LGC model to include a categorical latent class variable (Muthén, 2001; Muthén & Muthén, 2000; Muthén & Shedden, 1999). In contrast to a LGC model, latent growth mixture modeling (LGMM) allows the identification of qualitatively distinct growth trajectories of two or more latent classes (Bauer & Curran, 2003), and estimates the probability of membership into each latent class. It is typically assumed in LGMM that the functional form describing the overall change process in each latent class has no disjuncture. This assumption may become unrealistic for applications where repeated measures mixture data incorporates class-specific piecewise growth trajectory.

This leads to a need for expanding the LGMM to incorporate piecewise functions within each latent class. The objective of this study is to develop a piecewise linear-linear

LGMM, where the location of the knot is unknown in each latent class. This kind of model can be very useful for substantive researchers to address key questions such as in developmental and behavioral studies (e.g., substance abuse trajectories) or in studies seeking to measure the effect of some treatment or intervention (e.g., when individuals may need to seek professional services at the timing when mental ability decreases). The parameters of a piecewise linear-linear LGMM, including the knot, are estimated with ML via the expectation-maximization (EM) algorithm using existing statistical software. Estimation of piecewise linear-linear LGMM is carried out in *Mplus* 6.1 (Muthén & Muthén, 1998-2010), a popular SEM program.

The remainder of this document is organized as follows. Chapter 2 provides a review of the literature on LGC models, piecewise LGC models, and finite mixture models, including LGMMs. Chapter 3 describes the design of the proposed research, including methods of estimation and analysis. Chapter 4 provides the results, while Chapter 5 discusses the results and provides recommendations.

CHAPTER 2: REVIEW OF LITERATURE

This chapter reviews the literature on latent growth curve models, piecewise latent growth curve models, finite mixture models, and latent growth mixture models. It is important to review research on these models as all of these contribute significantly to the current body of knowledge about piecewise linear-linear latent growth mixture models where the estimation of the knot for each latent class is of primary interest.

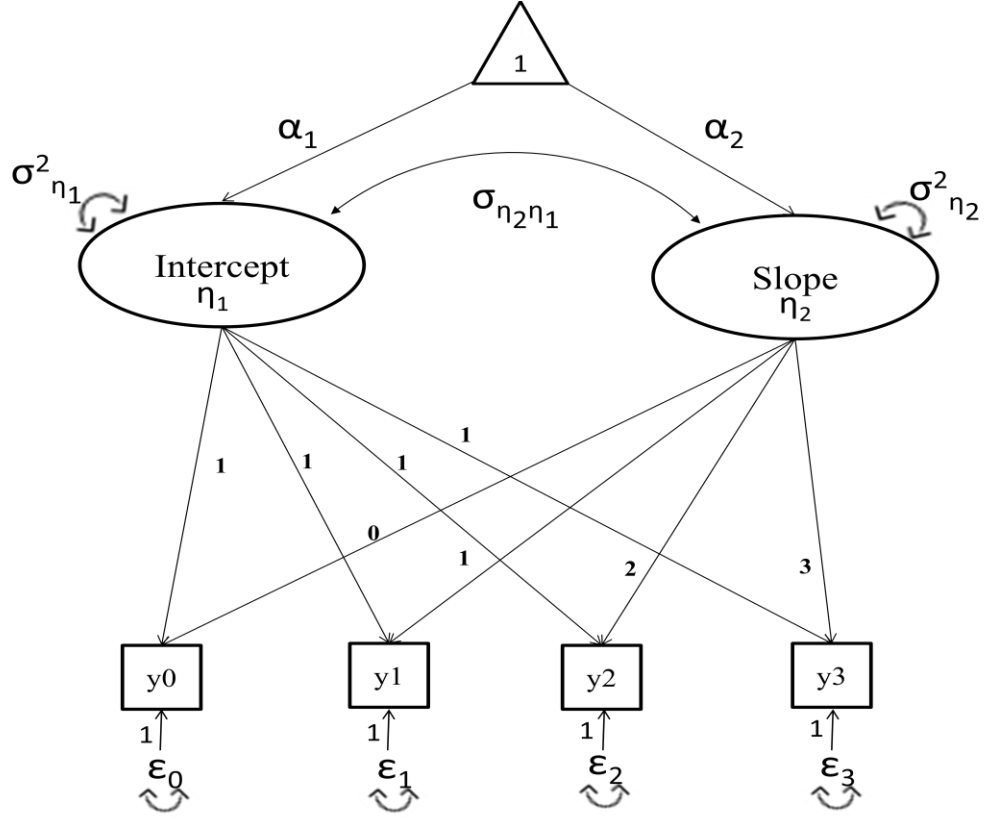
2.1 Latent Growth Curve Models

The LGC model was developed to analyze repeated measures data (Meredith & Tisak, 1990) where the form of the trajectory could be specified a priori or left as a partially parameterized model whose components could be estimated. The LGC model allows to disentangle the correlational structure of the repeated measures into intra-individual (within-person) variability as well as inter-individual (between-person) variability in individual subjects' growth characteristics across time (Preacher, Wichman, MacCallum, & Briggs, 2008). A typical application of LGC models specifies a function describing a linear change process often composed of two latent growth factors: (a) an intercept which describes initial level or status at some temporal reference point, and (b) a linear slope of growth which summarizes change over time. These two latent growth factors can be characterized by the mean value of intercept and slope and individual random variation and covariation around these two latent growth components (Duncan, Duncan, & Strycker, 2006).

For a fully specified model, the loadings from the intercept factor to each of the repeated measures are fixed to values of 1.0. That is, the intercept factor equally contributes to all repeated measures across all the waves of assessment. For the slope factor, the loadings are either fixed to describe linear change when theory specifies a linear form of growth or, alternatively, a LGC model with unspecified trajectory can be specified when theory dictates a single functional form of individual's growth but the functional form is not known a priori (Hancock & Lawrence, 2006; Meredith & Tisak, 1990). In this case, the pattern of loadings can be estimated as long as some of the loadings are fixed for model identification purposes and to set the per unit scale for growth. It is also possible to incorporate other functions in a LGC model that describe both linear and nonlinear change processes and a particular model may be chosen on theoretical grounds or via empirical exploration of the data. Furthermore, nonlinear models, including both higher-degree polynomials and intrinsically nonlinear functions, can be specified to capture important curvature in the response variable.

An LGC model that examines change across time in repeated measurements of an observed variable is termed a “first-order” LGC model. A path diagram of a conventional first-order linear LGC model with 4 equally spaced timepoints is provided in Figure 1.

Figure 1: A conventional first-order linear LGC model with linear trajectory and mean factor.



The basic formulation of a first-order LGC model includes two components: (1) a measurement model that connects the observed indicators with the corresponding latent growth factors across time, and (2) a structural model that describes the means and variances of latent growth factors.

Consider a set of repeated measures of a random variable Y for individual i , where the vector $\mathbf{y}_i = (y_1, y_2, \dots, y_{T_i})'$ is a set of T_i responses for individual i on Y . It is assumed that the distribution of \mathbf{y}_i is multivariate normal. The responses are observed on a set of repeated measurement occasions $\mathbf{t}_i = (1, \dots, T_i)'$, where T_i is the total number of observations for individual i . The subscript i on \mathbf{t}_i suggest that times of measurement

may vary from one person to another, which is very often seen in longitudinal data as a result of dropout, attrition, or by design of the longitudinal study. The general expression of measurement model for \mathbf{y}_i is (see, e.g., Blozis, 2004):

$$\mathbf{y}_i = \mathbf{\Lambda}_i \boldsymbol{\eta}_i + \boldsymbol{\epsilon}_i, \quad (2.11)$$

where $\mathbf{\Lambda}_i$ is a matrix of factor loadings reflecting the hypothesized underlying growth pattern:

$$\mathbf{\Lambda}_i = \begin{bmatrix} \lambda_{11} & \lambda_{12} & \cdots & \lambda_{1J} \\ \lambda_{21} & \lambda_{22} & \cdots & \lambda_{2J} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{T_i1} & \lambda_{T_i2} & \cdots & \lambda_{T_iJ} \end{bmatrix}.$$

The number of rows in $\mathbf{\Lambda}_i$ is equal to the number of measurement occasions on which individual i was observed. The columns of $\mathbf{\Lambda}_i$ have elements (e.g., λ_{11}) that define the shape of the growth curve over the observed measurement occasions. For example, in a linear LGC model, the factor loading matrix typically contains a column of ones for the intercept and a column of fixed values corresponding to increments of time (Willett & Sayer, 1994). The columns of $\mathbf{\Lambda}_i$ are commonly referred to as “basis” curves and are defined as the partial derivative of function f with respect to each of the model parameters. Furthermore, $\boldsymbol{\eta}_i = (\eta_{1i}, \eta_{2i}, \dots, \eta_{Ji})'$ is a vector of latent growth factors, and $\boldsymbol{\epsilon}_i = (\epsilon_1, \epsilon_2, \dots, \epsilon_{T_i})'$ is a vector of random errors or residuals that are often assumed to be normally distributed with mean zero and covariance matrix $\boldsymbol{\Theta}_i$ (i.e., $\boldsymbol{\epsilon}_i \sim N(0, \boldsymbol{\Theta}_i)$).

It is often assumed that the residuals are independent between measurement occasions with constant variance across time, once the linear dependence among the observed variables is accounted for. That is,

$$\boldsymbol{\Theta}_i = \text{diag}(\sigma_{\epsilon_i}^2, \sigma_{\epsilon_i}^2, \dots, \sigma_{\epsilon_i}^2),$$

where the $diag(.)$ operator refers to a diagonal matrix, and $\sigma_{\epsilon_i}^2$ refers to the variance of the residuals. It is, however, possible to specify different kinds of residual structures depending on other theoretical considerations or based on the characteristics of the data. A residual structure could accommodate the situation in which residuals are assumed to be independent between measurement occasions but the variances across time are allowed to be heterogeneous (Willett & Sayer, 1994). Of course, other residual covariance structures are possible but parameterization should be done as parsimoniously as possible.

The structural model is specified as (see, e.g., Bollen & Curran, 2006):

$$\boldsymbol{\eta}_i = \boldsymbol{\alpha} + \boldsymbol{\zeta}_i, \quad (2.12)$$

where $\boldsymbol{\alpha}$ is a $J \times 1$ vector of factor means, and $\boldsymbol{\zeta}_i$ is an $J \times 1$ is a vector of random disturbances in the first-order latent factors, $\boldsymbol{\eta}_i$, that are often assumed to be normally distributed with mean zero and covariance matrix $\boldsymbol{\Psi}$ (i.e., $\boldsymbol{\zeta}_i \sim N(0, \boldsymbol{\Psi})$). The covariance matrix $\boldsymbol{\Psi}$ is of the form:

$$\boldsymbol{\Psi} = \begin{bmatrix} \sigma_{\eta_1}^2 & & & \\ \sigma_{\eta_2\eta_1} & \sigma_{\eta_2}^2 & & \\ \vdots & \vdots & \ddots & \\ \sigma_{\eta_J\eta_1} & \sigma_{\eta_J\eta_2} & \dots & \sigma_{\eta_J}^2 \end{bmatrix},$$

where the diagonal elements are the variances of the growth factors and the off-diagonal elements are their covariances.

Like standard factor analysis models, the residuals are assumed to be uncorrelated with the continuous latent growth factors (i.e., $\text{cov}(\boldsymbol{\epsilon}_i, \boldsymbol{\eta}_i) = \mathbf{0}$). The residuals are also assumed to be uncorrelated with the random disturbances in the first-order latent factors (i.e., $\text{cov}(\boldsymbol{\epsilon}_i, \boldsymbol{\zeta}_i) = \mathbf{0}$). Furthermore, the residuals are assumed to be uncorrelated over

time (i.e., $\text{cov}(\epsilon_{im}, \zeta_{i,m+t}) = \mathbf{0}$ for $m \neq 0$). Given the preceding assumptions, the population mean vector and population covariance matrix of \mathbf{y}_i are (see, e.g., Blozis, 2004):

$$E(\mathbf{y}_i) = \boldsymbol{\mu}_i = \mathbf{\Lambda}_i \boldsymbol{\alpha} \quad (2.13)$$

and

$$\text{cov}(\mathbf{y}_i) = \boldsymbol{\Sigma}_i = \mathbf{\Lambda}_i \boldsymbol{\Psi} \mathbf{\Lambda}_i' + \boldsymbol{\Theta}_i. \quad (2.14)$$

2.1.1 ML estimation of latent growth curve models

The parameters of the LGC model can be denoted as $\boldsymbol{\Omega}$, a $P \times 1$ vector, which consists of all free parameters in $\boldsymbol{\alpha}$, $\boldsymbol{\Psi}$, and $\boldsymbol{\Theta}_i$ (i.e., $\boldsymbol{\Omega} = (\boldsymbol{\alpha}', \text{vech}(\boldsymbol{\Psi})', \text{vech}(\boldsymbol{\Theta}_i)')'$), where the $\text{vech}(\cdot)$ operator creates a column vector of symmetric matrices ($(\boldsymbol{\Psi})$ and $(\boldsymbol{\Theta}_i)$) by stacking successive row-wise elements of the lower triangle below one another. The parameters $\boldsymbol{\Omega}$ can be estimated via ML. The ML estimator is the most popular method of estimation to use with LGC models because it has a number of desirable properties, such as the estimated model parameters are consistent, asymptotically unbiased, asymptotically normal, and asymptotically efficient (Bollen & Curran, 2006).

Let $\boldsymbol{\mu}_i(\boldsymbol{\Omega})$ denote the model-implied mean vector and $\boldsymbol{\Sigma}_i(\boldsymbol{\Omega})$ denote the model-implied covariance matrix. Both $\boldsymbol{\mu}_i(\boldsymbol{\Omega})$ and $\boldsymbol{\Sigma}_i(\boldsymbol{\Omega})$ are functions of the parameters of the LGC model. Because the population mean vector $\boldsymbol{\mu}$ and population covariance matrix $\boldsymbol{\Sigma}$ are unavailable, the sample mean vector $\bar{\mathbf{y}}_i$ and sample covariance matrix \mathbf{S}_i are used as estimates. The goal is to choose values of the estimated model parameters $\hat{\boldsymbol{\Omega}}$ (i.e., $\hat{\boldsymbol{\alpha}}$, $\hat{\boldsymbol{\Psi}}$, and $\hat{\boldsymbol{\Theta}}_i$) such that $\boldsymbol{\mu}_i(\hat{\boldsymbol{\Omega}})$ is close to $\bar{\mathbf{y}}_i$ and $\boldsymbol{\Sigma}_i(\hat{\boldsymbol{\Omega}})$ is close to \mathbf{S}_i (Bollen & Curran, 2006).

In typical ML estimation, the log-likelihood function is maximized with respect to model parameters to obtain estimates. That is (see, e.g., Preacher et al., 2008),

$$\log L = -\frac{1}{2} \sum_{i=1}^n \{T_i \log 2\pi + \log |\Sigma_i(\Omega)| + [\mathbf{y}_i - \boldsymbol{\mu}_i(\Omega)]' \Sigma_i^{-1}(\Omega) [\mathbf{y}_i - \boldsymbol{\mu}_i(\Omega)]\}. \quad (2.15)$$

In a typical SEM framework, however, $\log L$ is maximized when the following *discrepancy function* (F_{ML}) is minimized (see, e.g., Bollen & Curran, 2006; Preacher et al., 2008):

$$\begin{aligned} F_{ML} = & \log |\Sigma_i(\Omega)| - \log |\mathbf{S}_i| + \text{tr}[\{\mathbf{S}_i - \Sigma_i(\Omega)\} \Sigma_i^{-1}(\Omega)] \\ & + [\bar{\mathbf{y}}_i - \boldsymbol{\mu}_i(\Omega)]' \Sigma_i^{-1}(\Omega) [\bar{\mathbf{y}}_i - \boldsymbol{\mu}_i(\Omega)]. \end{aligned} \quad (2.16)$$

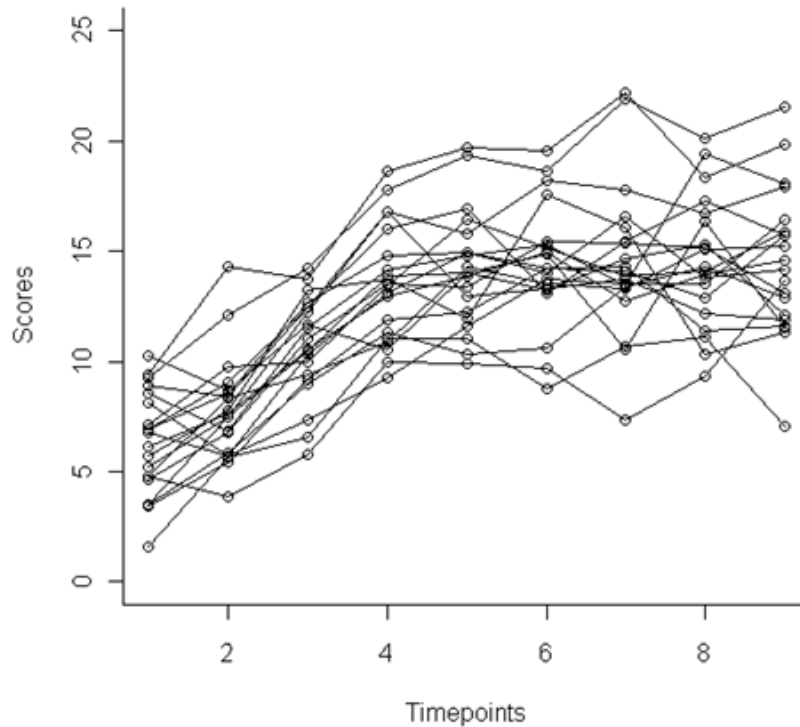
Thus, the idea behind ML estimation is to compute $\hat{\Omega}$, given sample based $\bar{\mathbf{y}}_i$ and \mathbf{S}_i , so that it produces model-implied $\boldsymbol{\mu}_i(\hat{\Omega})$ and $\Sigma_i(\hat{\Omega})$ matrices that minimize F_{ML} . Note that minimization of the discrepancy function F_{ML} assumes that complete sample data are used to obtain $\bar{\mathbf{y}}_i$ and \mathbf{S}_i (Preacher et al., 2008). If the data are incomplete and missing at random, *full information maximum likelihood* (FIML) can be used to obtain ML parameter estimates. In this case, FIML estimation allows the LGC model to be fit directly to raw incomplete data (see e.g., Enders, 2006).

2.2 Piecewise Latent Growth Curve Models

It is typically assumed in LGC models that the functional form describing the overall change process in the repeated measures data is continuous. However, assuming a single uninterrupted functional form underling the overall change process may be improbable for applications where data are comprised of different growth phases. Figure

2 provides a graphical representation of a generic LGC model that comprises of two different linear-linear growth phases.

Figure 2: A spaghetti plot depicting generic piecewise linear-linear change process.



Piecewise LGC models, an extension of LGC models, are flexible because each phase can be specified to conform to a particular functional form of the overall change process (Cudeck & Harring, 2010). The term “piecewise” is obtained from the piecewise regression model, which is a special case of spline regression model (Marsh & Cormier, 2001). To make more concrete, consider a linear-linear piecewise process. In this situation, the formulated model assumes a simple regression line for the dependent variable, but with possibly different parameterizations in different ranges of the predictor (Bates & Watts, 1988; see also Seber & Wild, 1989, Ch. 9). Note that the assumptions

underlying a piecewise LGC model are similar to the assumptions underlying LGC model discussed in the previous section.

One of the more interesting features of a piecewise model is the knot. The knot is the value of the predictor at which the function shifts from the first phase to the other (Cudeck, 1996; Cudeck & Klebe, 2002). The value of knot can be either fixed or estimated. Consider a two-phase piecewise linear-linear LGC model, for example, with a *known* value of the knot (see, e.g., Chou et al., 2004):

$$y_{ji} = D_1(\eta_{1i} + \eta_{2i}t_{ji}) + D_2(\eta_{3i} + \eta_{4i}t_{ji}) + \epsilon_{ji}, \quad (2.21)$$

where y_{ji} is the observed response of individual i at time j ; t_{ji} represents the time of measurement; η_{1i} and η_{2i} are the intercept and slope growth factors of the first phase, respectively; η_{3i} and η_{4i} are the intercept and slope growth factors of the second phase, respectively; ϵ_{ji} is the random normal error; and D_1 and D_2 are the two dummy-coded variables that take the values of either 0 or 1, depending on the phase from which y_{ji} was obtained. That is, $D_1 = 1$ in the first phase and $D_1 = 0$ elsewhere. Similarly, $D_2 = 1$ in the second phase and $D_2 = 0$ elsewhere. In Equation 2.21 there are four linear coefficients (i.e., growth factors), $\boldsymbol{\beta}_i = (\eta_{1i}, \eta_{2i}, \eta_{3i}, \eta_{4i})'$. One of the coefficients, however, can be eliminated because it is often assumed in piecewise models that the two separate functions join at the knot, γ . That is, $\eta_{1i} + \eta_{2i}\gamma = \eta_{3i} + \eta_{4i}\gamma$, and thus one coefficient is redundant. The decision as to which coefficient to eliminate is arbitrary, unless the researcher has a theory for eliminating a specific coefficient. Following the above alternative, one may set the intercept of the second phase as:

$$\eta_{3i} = \eta_{1i} + \eta_{2i}\gamma - \eta_{4i}\gamma.$$

The expression for η_{3i} can then be substituted into Equation 2.21 as:

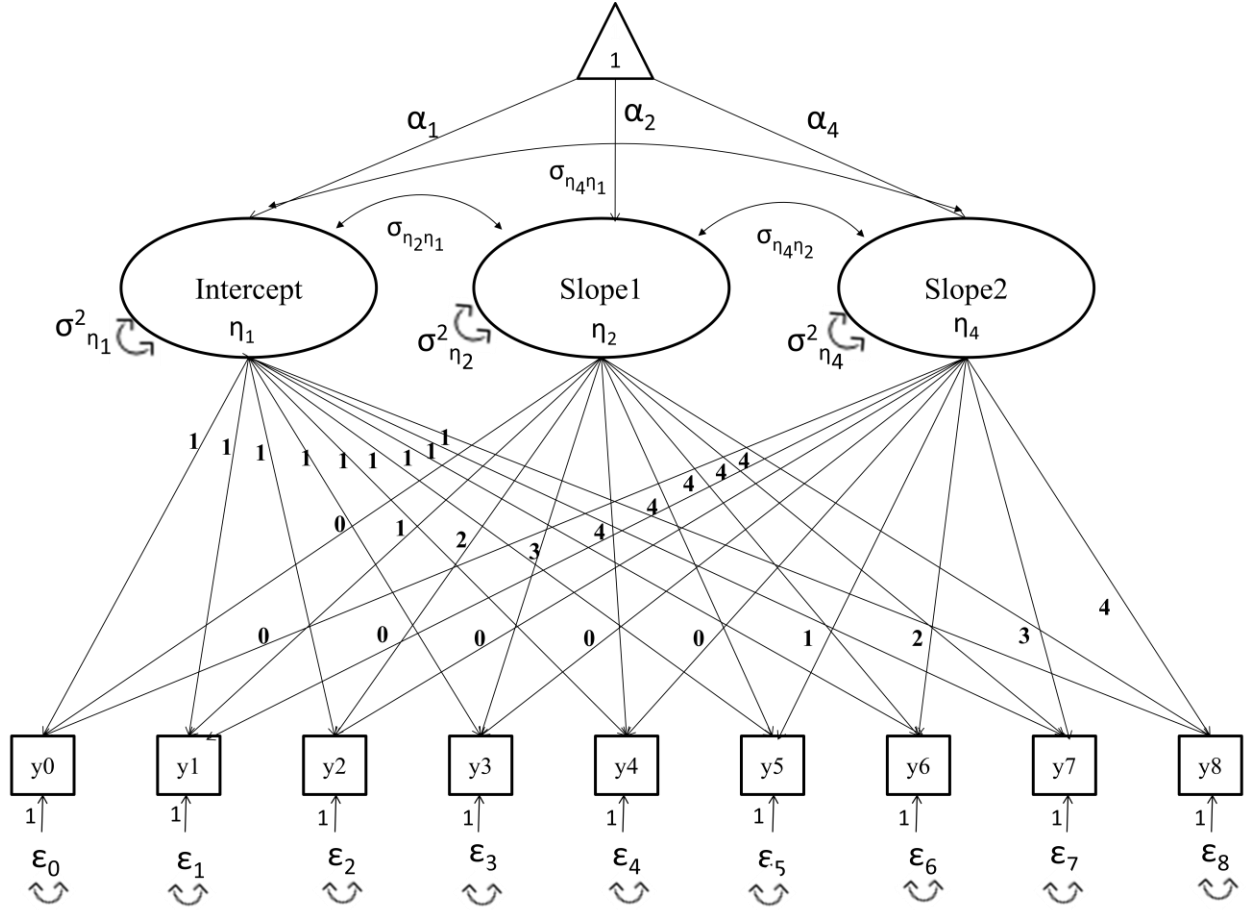
$$y_{ji} = D_1(\eta_{1i} + \eta_{2i}t_{ji}) + D_2(\eta_{1i} + \eta_{2i}\gamma + \eta_{4i}(t_{ji} - \gamma)) + \epsilon_{ji}. \quad (2.22)$$

Note that in Equation 2.22 the value of knot is known a priori, thus there are three freely estimable coefficients, $\boldsymbol{\beta}_i = (\eta_{1i}, \eta_{2i}, \eta_{4i})'$. Furthermore, the two-phase piecewise linear-linear LGC model with a *known* value of the knot fits into the general model of Equation 2.11 by coding the j th row of the factor loading matrix $\boldsymbol{\Lambda}_i$ according to whether t_{ji} is greater than the knot . That is (see, e.g., Harring et al., 2006),

$$\boldsymbol{\Lambda}_i = \begin{cases} [1, & t_{ji}, & 0] & t_{ji} \leq \gamma \\ [1, & \gamma, & (t_{ji} - \gamma)] & t_{ji} > \gamma \end{cases} \quad (2.23)$$

where the columns of $\boldsymbol{\Lambda}_i$ are the partial derivative of the function described in Equation 2.22 with respect to each of the model parameters. A path diagram of a conventional piecewise linear-linear LGC model with 9 equally spaced timepoints where the location of knot is *known* is provided in Figure 3.

Figure 3: A conventional piecewise linear-linear LGC model where the location of knot is known.



Note that in the model depicted in Figure 3, $\mathbf{t}_i = (0, 1, 2, 3, 4, 5, 6, 7, 8)'$ and the location of the knot is $\gamma = 4$. Hence, following Equation 2.23 the factor loading matrix $\mathbf{\Lambda}_i$ of the model shown in Figure 3 is:

$$\mathbf{\Lambda}_i = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 2 & 0 \\ 1 & 3 & 0 \\ 1 & 4 & 0 \\ 1 & 4 & 1 \\ 1 & 4 & 2 \\ 1 & 4 & 3 \\ 1 & 4 & 4 \end{bmatrix}$$

Alternatively, instead of assuming the value of the knot to be a fixed known value, one can estimate the knot by specifying it as one of the parameters in the model. Harring, Cudeck, and du Toit (2006) developed a first-order piecewise LGC model where the location of knot was unknown for investigating individual behavior that exhibited distinct phases of development in observed variables only. The specification of a two-phase piecewise linear-linear LGC model, for example, with an *unknown* value of the knot is (see, e.g., Harring et al., 2006):

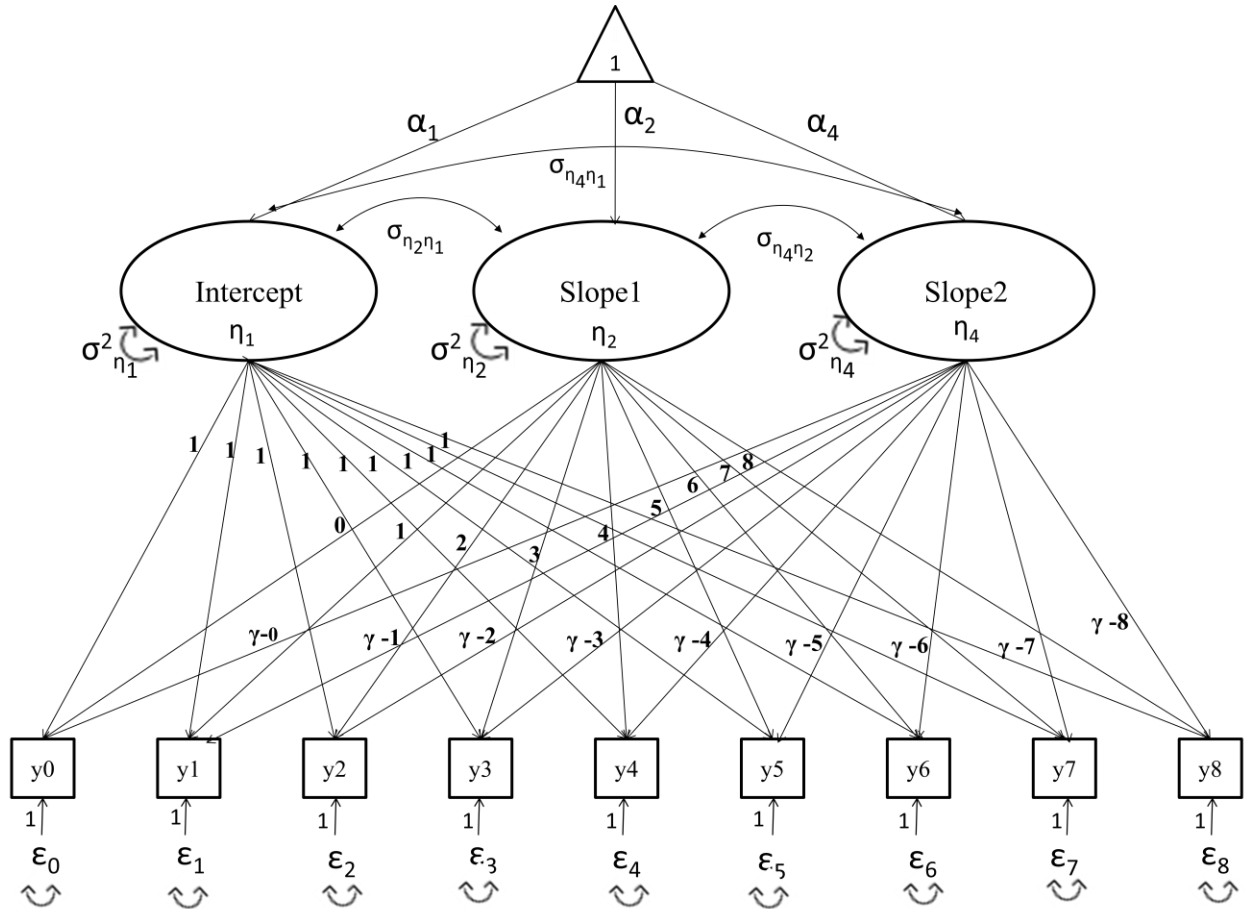
$$y_{ji} = \begin{cases} \eta_{1i} + \eta_{2i}t_{ji} + \epsilon_{ji} & t_{ji} \leq \gamma \\ \eta_{3i} + \eta_{4i}t_{ji} + \epsilon_{ji} & t_{ji} > \gamma \end{cases}, \quad (2.24)$$

where y_{ji} is the observed response of individual i at time j ; t_{ji} represents the time of measurement; η_{1i} and η_{2i} represent the intercept and slope growth factors for the first phase, respectively; η_{3i} and η_{4i} represent the intercept and slope growth factors for the second phase, respectively; and γ represents the knot. In Equation 2.24 there are four linear coefficients, $\boldsymbol{\beta}_i = (\eta_{1i}, \eta_{2i}, \eta_{3i}, \eta_{4i})'$ and one nonlinear knot γ . It is often assumed in piecewise models that the two separate functions join at the knot; hence, there are effectively three free coefficients and one nonlinear knot in the target function, that is, $(\eta_{1i}, \eta_{2i}, \eta_{4i}, \gamma)'$. Note that in Equation 2.24 there is no ' i ' subscript for the knot γ . This is because in this model it is assumed that the location of unknown knot is fixed, hence, there is no variability around the location of unknown knot.

Furthermore, the parameterization of the model in Equation 2.24 cannot be specified directly in a SEM framework in a way that permits the estimation of γ . The difficulty stems from the inability of the existing SEM software packages to incorporate executable programming functions, like if-then statements, in the estimation step. A common alternative is to reparameterize the piecewise LGC model so that it fits within

the system of SEM software packages. The issues related to the estimation of the piecewise LGC models are discussed in more detail in a later section. A path diagram of a conventional piecewise linear-linear LGC model with 9 equally spaced timepoints where the location of knot is *unknown* is provided in Figure 4.

Figure 4: A conventional piecewise linear-linear LGC model where the location of knot is unknown.



Note that the specification of the factor loading matrix $\mathbf{\Lambda}_i$ of the reparameterized piecewise linear-linear LGC model depicted in Figure 4 is similar to the specification in Harring et al. (2006). That is,

$$\mathbf{\Lambda}_i = \begin{bmatrix} 1 & 0 & \sqrt{(\gamma - 0)^2} \\ 1 & 1 & \sqrt{(\gamma - 1)^2} \\ 1 & 2 & \sqrt{(\gamma - 2)^2} \\ 1 & 3 & \sqrt{(\gamma - 3)^2} \\ 1 & 4 & \sqrt{(\gamma - 4)^2} \\ 1 & 5 & \sqrt{(\gamma - 5)^2} \\ 1 & 6 & \sqrt{(\gamma - 6)^2} \\ 1 & 7 & \sqrt{(\gamma - 7)^2} \\ 1 & 8 & \sqrt{(\gamma - 8)^2} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \gamma - 0 \\ 1 & 1 & \gamma - 1 \\ 1 & 2 & \gamma - 2 \\ 1 & 3 & \gamma - 3 \\ 1 & 4 & \gamma - 4 \\ 1 & 5 & \gamma - 5 \\ 1 & 6 & \gamma - 6 \\ 1 & 7 & \gamma - 7 \\ 1 & 8 & \gamma - 8 \end{bmatrix}$$

Of course, this modeling framework provides sufficient flexibility to summarize other functional forms in the different phases with the stipulation that each phase does not have to conform to the same function. For example, a two-phase piecewise quadratic-linear LGC model with *unknown* knot could be specified when the trajectory in the first developmental stage has some curvature and the trajectory in the second developmental stage is straight-line. This kind of model can be easily specified by extending the linear-linear piecewise function (see, e.g., Cudeck & Harring, 2010). For example,

$$y_{ji} = \begin{cases} \eta_{1i} + \eta_{2i}t_{ji} + \eta_{3i}t_{ji}^2 + \epsilon_{ji} & t_{ji} \leq \gamma \\ \eta_{4i} + \eta_{5i}t_{ji} + \epsilon_{ji} & t_{ji} > \gamma \end{cases}, \quad (2.25)$$

where the first phase of model corresponds to a quadratic function and the second phase of model corresponds to a linear function. Note that two restrictions can be imposed on Equation (2.25). The first restriction is: $\eta_{1i} + \eta_{2i}\gamma + \eta_{3i}\gamma^2 = \eta_{4i} + \eta_{5i}\gamma$. With this restriction in place, a parameter in Equation (2.25) can be eliminated as it is redundant. The second restriction that can be imposed on the model is: $\eta_{2i} + 2\eta_{3i}\gamma = \eta_{5i}$, which allows the two functions to meet at the knot in a smooth transition. As a result of this second restriction, the linear coefficient of the first segment can be expressed in terms of

the others. It is thus possible to specify different kinds of piecewise LGC models depending on the characteristics underlying the data.

2.2.1 ML estimation of piecewise linear-linear latent growth curve models with unknown knot

There are three linear coefficients, $\beta_i = (\eta_{1i}, \eta_{2i}, \eta_{4i})'$, and one nonlinear coefficient, γ , in a piecewise linear-linear LGC model. It is, however, not possible to specify the parameterization in Equation 2.24 directly in an SEM framework in a way that permits treating γ , the knot, as an estimated parameter. A common alternative tactic is to reparameterize the piecewise LGC model. Reparameterized piecewise LGC models have the same number of free parameters as in a linear-linear piecewise LGC model, but are fit within the system employed by many SEM software packages. Additionally, reparameterization of the model makes it convenient for estimation to be carried out by ML estimation in a typical SEM framework (Harring et al., 2006). Upon convergence, the estimated parameters of the reparameterized model are then transformed back to the original parameters of the piecewise linear-linear LGC model. The only limitation of reparameterization is that the fit of the model may be affected by the transformation from one version of a model into another form. Harring et al. (2006) mentioned that generally the difference in fit is not great, and any slight loss in fit would seem to be offset by the ease with which the reparameterized model can be estimated. A detailed description of reparameterization procedure is discussed in Appendix A. The procedure of transformation of estimated parameters of reparameterized model back to the parameters of the original model is discussed in Appendix B.

2.3 Finite Mixture Models

In many research settings the observed sample can be seen as stemming from multiple population distributions with distinct characteristics. A composite of two or more subpopulation distributions is called a finite mixture distribution (see, e.g., McLachlan & Peel, 2000) and the subpopulations are called mixture components, where component membership is unobserved for each observation. It is of interest to build a generic model that allows us to combine the distributions from different subpopulations. A finite mixture model is a probability model that combines the probability densities across all the subpopulations underlying the data. The general form of a finite mixture model is given as:

$$f(\mathbf{y}) = \varphi_1 f_1(\mathbf{y}) + \cdots + \varphi_K f_K(\mathbf{y}), \quad (2.31)$$

where $f(\mathbf{y})$ is the composite density function for all $k = 1, \dots, K$ number of components. A single density $f_k(\mathbf{y})$ is referred as the component density. It is typically assumed that the distribution underlying the subpopulations (mixture components) have the same density. The parameters $\varphi_1, \dots, \varphi_K$ are called the mixing proportions. It is assumed that the mixing proportions φ_k are non-negative quantities that sum to one. That is,

$$\sum_{k=1}^K \varphi_k = 1, \text{ where } \varphi_k \geq 0.$$

In a finite mixture model each component distribution has its own set of parameters denoted as $\boldsymbol{\theta}_k$. Typically, $\boldsymbol{\theta}_k$ are unknown parameter values that must be estimated from sample data. This is also the case with the number of components. That is, we often do not know how many components are in the model, thus we have to infer the optimal number of components from the sample data.

There are two main purposes for using finite mixture models. The first purpose is to provide a framework to approximate complex distributions with two or more component distributions. The second purpose is to use finite mixture models as a model-based clustering tool that can help to identify more than one unobserved population with the intent to infer qualitatively distinct classes of individuals in the population. When modeling population heterogeneity using finite mixture models, it is typically assumed that data came from a mixture of two or more distributions from the same parametric family with parameters that are allowed to differ across components (Frühwirth-Schnatter, 2006; McLachlan & Peel, 2000).

2.3.1 Mixtures of univariate distributions

The probability density function of a two-class mixture of univariate distributions, for example normal, is specified as:

$$f(\mathbf{y}) = \varphi_1 f_1(\mathbf{y}) + (1 - \varphi_1) f_2(\mathbf{y}), \quad (2.32)$$

where $f(\mathbf{y})$ is the composite density function; $f_1(\mathbf{y})$ and $f_2(\mathbf{y})$ are the component densities of the two classes; and φ_1 and $(1 - \varphi_1)$ are the mixing proportions. The component densities in this example are defined to be normally distributed so that

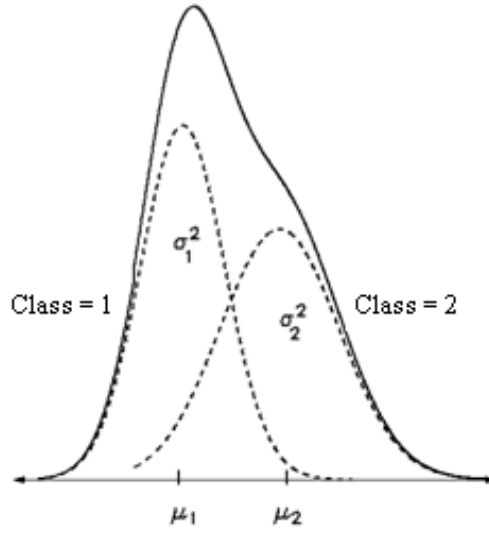
$$f_1(\mathbf{y}) = \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp \left[-\frac{(\mathbf{y} - (\mathbf{1} \cdot \mu_1))^2}{2\sigma_1^2} \right]$$

$$f_2(\mathbf{y}) = \frac{1}{\sqrt{2\pi\sigma_2^2}} \exp \left[-\frac{(\mathbf{y} - (\mathbf{1} \cdot \mu_2))^2}{2\sigma_2^2} \right],$$

where \mathbf{y} denotes the value of observations on variable Y ; $\mathbf{1}$ denotes the unity vector; the dot (\cdot) denotes the multiplication sign; μ_1 and σ_1^2 are the class mean and class variance of population 1; and μ_2 and σ_2^2 are the class mean and class variance of population 2. Note

that the density function for each component in Equation 2.32 is described by two parameters, the class mean and class variance. To fit the model in Equation 2.32, it is thus necessary to estimate the parameters $(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2, \varphi_1)$. A graphical representation of a two-class univariate normal mixture model is shown in Figure 5.

Figure 5: A two-class univariate normal mixture model.



Mixtures of regression models, also known as latent class regression models, are used to capture parameter heterogeneity for cross-sectional data (Frühwirth-Schnatter, 2006). Unlike typical regression analysis, which assumes that the distribution of data is governed by one set of parameters, simple linear regression mixture models allow for different sets of parameters, each corresponding to an underlying latent class (Grün & Leisch, 2006). Individuals within each latent class share the same regression function. A finite mixture of regression models has a class-specific probability density function (pdf).

$$f_k(\mathbf{y}|\boldsymbol{\beta}'_k\mathbf{x}, \sigma_k^2) = \frac{1}{\sqrt{2\pi\sigma_k^2}} \exp\left[-\frac{(\mathbf{y} - \boldsymbol{\beta}'_k\mathbf{x})'(\mathbf{y} - \boldsymbol{\beta}'_k\mathbf{x})}{2\sigma_k^2}\right], \quad (2.33)$$

with class-specific residual variance σ_k^2 , and vector of class-specific regression coefficients $\boldsymbol{\beta}_k$.

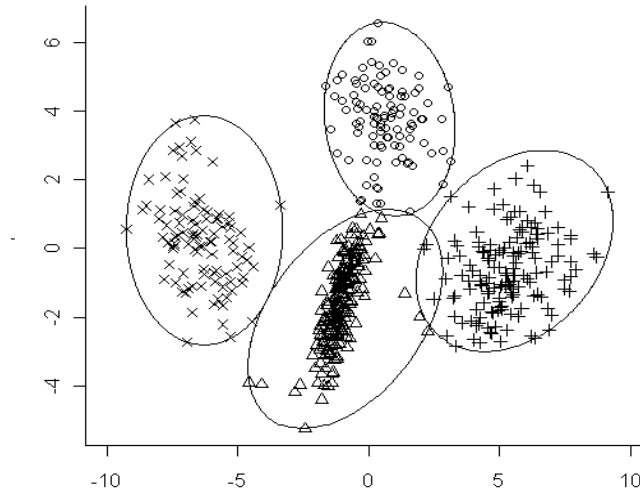
2.3.2 Mixtures of multivariate distributions

Mixtures of multivariate distributions such as multivariate normal, with a class-specific probability density function (pdf), are written as

$$f_k(\mathbf{y}_i | \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) = \frac{1}{(2\pi)^{P/2} |\boldsymbol{\Sigma}_k|^{1/2}} \exp \left[-\frac{1}{2} (\mathbf{y}_i - \boldsymbol{\mu}_k)' \boldsymbol{\Sigma}_k^{-1} (\mathbf{y}_i - \boldsymbol{\mu}_k) \right], \quad (2.34)$$

where \mathbf{y}_i denote a P-dimensional vector containing the scores for individual i on a set of P observed continuous random variables, $\boldsymbol{\mu}_k$ is a vector of class-specific means, and $\boldsymbol{\Sigma}_k$ is a class-specific variance-covariance matrix. A graphical representation of a multivariate normal mixture (specifically, bivariate) is shown in Figure 6.

Figure 6: A multivariate normal mixture model.



2.3.3 Latent growth mixture models

LGMMs are a kind of multivariate normal mixture model. That is, both multivariate normal mixture models and latent growth mixture models assume that the continuous observed data in \mathbf{y} are a mixture of two or more unobserved subpopulation

distributions where \mathbf{y} is assumed to be multivariate normally distributed. The main difference between the two models is that in multivariate normal mixture models the mean vector and covariance matrix are unstructured, whereas in LGMMs the mean vector and covariance matrix have specific structures related to a hypothesized growth form.

An alternative explanation of LGMM is that it is a statistical technique that combines LGC modeling with the idea of latent classes (Muthén, 2001; 2002; Muthén & Muthén, 2000; Muthén & Shedden, 1999). Contrary to the LGC model, which assumes that all individuals come from a single population and share the same growth pattern, LGMM relaxes the single population assumption by allowing the observed data to come from a mixture of two or more unobserved subpopulations (i.e., latent classes). LGMM allows the identification of latent classes, if they exist, that follow qualitatively distinct growth trajectories (Bauer & Curran, 2003). This is accomplished by using a categorical latent variable to represent two or more distinct trajectory classes. The combined use of continuous and categorical latent variables allows individuals to vary around the mean growth curve for their particular subgroup where each subgroup has its own model parameter values (Bauer & Curran, 2003; Muthén, 2001; Muthén & Shedden, 1999).

Suppose the observed data come from K subpopulations ($k = 1, \dots, K$), with a latent categorical variable indicating the latent class membership for individual i , where k designates each latent class and indicates that model parameter values may differ across classes. Assuming conditional independence, the class-specific measurement portion of the model is specified as (see, e.g., Muthén & Shedden, 1999):

$$\mathbf{y}_i = \mathbf{\Lambda}_{ik}\boldsymbol{\eta}_i + \boldsymbol{\epsilon}_i, \text{ for } k = 1, \dots, K, \quad (2.35)$$

where

$$\boldsymbol{\epsilon}_i \sim N(0, \boldsymbol{\Theta}_{ik}).$$

The specification of Equation 2.35 is similar to that given for conventional LGC model specified in Equation 2.11. That is, the vector \mathbf{y}_i is a set of responses for individual i on a set of repeated measurement occasions $\mathbf{t}_i = (1, \dots, T_i)'$, where T_i is the total number of observations for the individual i ; $\boldsymbol{\Lambda}_{ik}$ is a matrix of factor loadings; $\boldsymbol{\eta}_i$ is a vector of continuous latent growth factors particular to individual i ; and finally, $\boldsymbol{\epsilon}_i$ is a vector of residuals. Note that $\boldsymbol{\Theta}_{ik}$ is typically assumed to be diagonal,

$$\boldsymbol{\Theta}_{ik} = \text{diag}(\sigma_{\epsilon_k}^2, \sigma_{\epsilon_k}^2, \dots, \sigma_{\epsilon_k}^2)',$$

where the $\text{diag}(\cdot)$ operator creates a diagonal matrix; and $\sigma_{\epsilon_k}^2$ refers to the variance of the measurement residuals.

The class-specific structural component of the model is specified as

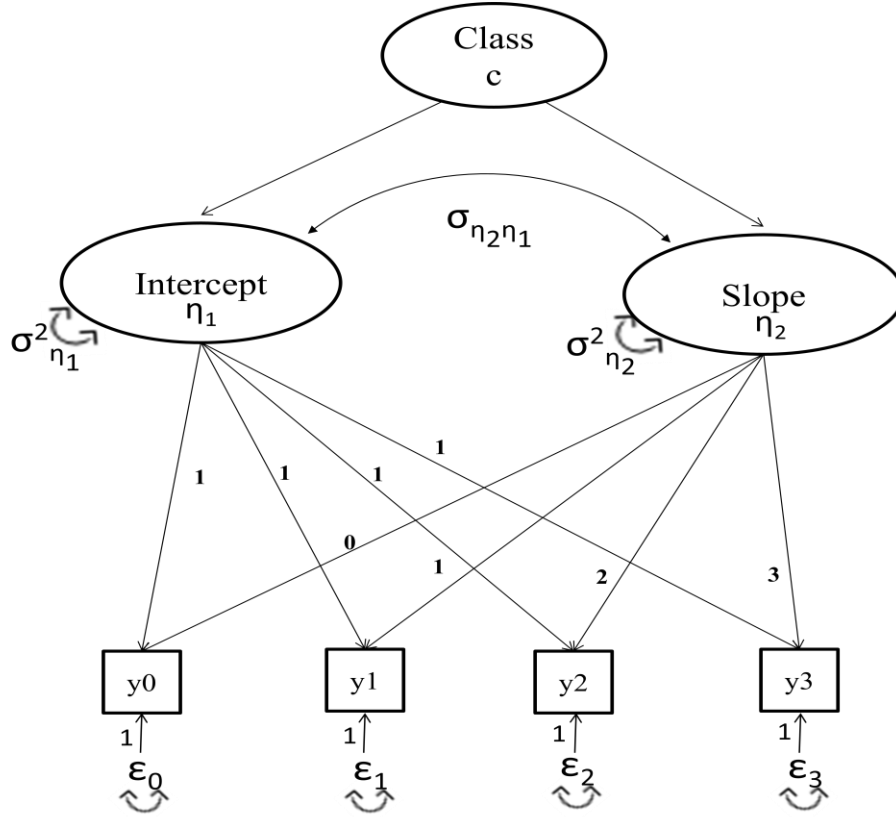
$$\boldsymbol{\eta}_i = \boldsymbol{\alpha}_k + \boldsymbol{\zeta}_i, \quad (2.36)$$

where

$$\boldsymbol{\zeta}_i \sim N(0, \boldsymbol{\Psi}_k), \text{ for } k = 1, \dots, K.$$

The specification of Equation 2.36 is similar to that given for conventional LGC model specified in Equation 2.12. That is $\boldsymbol{\alpha}_k$ is a vector of factor means and $\boldsymbol{\zeta}_i$ is a vector of random disturbances in the first-order latent factors, $\boldsymbol{\eta}_i$. Note that the subscript k in Equations 2.35 and 2.36 indicates a separate model for each latent class k , thus allowing for heterogeneity within the population. A path diagram of a conventional two class linear LGMM with four equally spaced timepoints is provided in Figure 7.

Figure 7: A conventional two-class linear LGMM.



Just like in finite mixture modeling, where it is assumed that group membership is unobserved and must be estimated along with the other parameters of the model, in LGMM it is assumed that the proportion φ_k of cases falling in latent class $k = 1, \dots, K$ is unknown and must be estimated. Thus, proportions φ_k are parameters to be estimated in addition to those parameters found in the standard LGC model. Additionally, proportions φ_k are non-negative quantities that sum to one, hence the number of these free parameters is $K-1$. That is,

$$\sum_{k=1}^K \varphi_k = 1, \text{ where } \varphi_k \geq 0.$$

2.3.4 ML estimation of latent growth mixture models via the EM Algorithm

The parameters of a LGMM, $\boldsymbol{\Omega} = (\boldsymbol{\theta}'_1, \dots, \boldsymbol{\theta}'_K, \varphi_1, \dots, \varphi_{K-1})'$ can be estimated by using ML estimation via the EM algorithm (Muthén & Shedden, 1999). Note that when estimating parameters in LGMM, the parameters related to the LGC model, along with the class proportions φ_k , are estimated. The log-likelihood function of observed data \mathbf{y}_i is given as (see, e.g., Muthén & Shedden, 1999; Tolvanen, 2008):

$$\log L = \log \prod_{i=1}^n L_i = \sum_{i=1}^n \log L_i = \sum_{i=1}^n f(\mathbf{y}_i), \quad (2.37)$$

where density function f is mixed from K density functions

$$f(\mathbf{y}) = \varphi_1 f_1(\mathbf{y}) + \dots + \varphi_K f_K(\mathbf{y}).$$

The density function for class k is

$$f_k(\mathbf{y}) \sim N(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$$

where

$$\boldsymbol{\mu}_k = \boldsymbol{\Lambda}_{ik} \boldsymbol{\alpha}_k$$

$$\boldsymbol{\Sigma}_k = \boldsymbol{\Lambda}_{ik} \boldsymbol{\Psi}_k \boldsymbol{\Lambda}_{ik}' + \boldsymbol{\Theta}_{ik}.$$

In typical ML estimation, the log-likelihood is maximized with respect to model parameters to obtain estimates. However, because LGMMs contain latent variable values and latent class memberships that are both unobserved, there is no closed-form solution for the parameter estimates (e.g., Mann, 2009). Thus, the EM algorithm is needed to obtain the model parameter estimates.

The EM algorithm obtains ML parameter estimates in the presence of missing data. In the context of LGMM, the missing part is denoted as the class information vector $\mathbf{c}_i = (c_{i1}, \dots, c_{iK})'$, where $c_{ik} = 1$ if y_i was produced by the k th component, otherwise

$c_{ik} = 0$. The complete-data log-likelihood (i.e., if complete data vector $\mathbf{x} = (\mathbf{y}_i', \mathbf{c}_i')$ was observed) is

$$L(\boldsymbol{\Omega}, \mathbf{y}_i, \mathbf{c}_i) = \sum_{i=1}^n \left[\sum_{k=1}^K c_{ik} \log p(c_{ik} = 1) + \sum_{k=1}^K c_{ik} \log f(\mathbf{y}_i | c_{ik}) \right]. \quad (2.38)$$

In the E-step of the EM algorithm, the conditional expectation of the probability of class membership is computed given the observed data \mathbf{y}_i and the current parameter estimate $\hat{\boldsymbol{\Omega}}_0$ (the initial starting value for $\boldsymbol{\Omega}$ on the first iteration) and the values from M-step in further iterations. The posterior probabilities (from Bayes' theorem) for observation i belonging to class k is calculated in the E step using formula:

$$\varphi_k = \frac{\varphi_K f_k(\mathbf{y}_i)}{f(\mathbf{y}_i)}, i = 1, \dots, n \text{ and } k = 1, \dots, K. \quad (2.39)$$

These posterior probabilities are then used in M-step when maximizing expected values in Equation 2.38:

$$E \left[\sum_{i=1}^n \sum_{k=1}^K c_{ik} \log p(c_{ik} = 1) \right] = \sum_{i=1}^n \sum_{k=1}^K \varphi_k \log p(c_{ik} = 1), \quad (2.40)$$

resulting in the φ_k parameters in the Equation 2.38 and to maximize

$$\sum_{i=1}^n \sum_{k=1}^K \varphi_k \log f_k(\mathbf{y}_i).$$

resulting in $f_k(\mathbf{y})$ in Equation 2.38 with the estimates of $\boldsymbol{\Lambda}_{ik}$, $\boldsymbol{\alpha}_k$, $\boldsymbol{\Psi}_k$, and $\boldsymbol{\Theta}_{ik}$.

After the M-step, the algorithm returns to the E-step to calculate new posterior probabilities and then again to the M-step. This iteration continues until the convergence criterion related to the complete-data log-likelihood is met. Note that Mplus 6 (Muthén & Muthén, 1998-2010) uses ML estimation via the EM algorithm; a description of the

estimation in general for latent variable mixture modeling is provided in the *Mplus* 6 technical appendices (Muthén & Muthén, 1998-2010).

2.4 Research Goals

There are two main objectives of this dissertation research. The first objective is to extend the framework of LGMM to a two-class piecewise linear-linear LGMM where the location of the knot in each latent class is unknown. The basic idea is to combine the piecewise linear-linear LGC model with latent classes. That is, each latent class has its own qualitatively distinct piecewise growth trajectory. To accomplish this objective, a series of Monte Carlo simulations empirically investigate the ability of two-class piecewise linear-linear LGMMs to recover true (known) growth parameters of distinct populations under different manipulated conditions. Specifically, the current research compares the performance of the two-class piecewise linear-linear LGMM under different manipulated conditions of 1) sample size, 2) class mixing proportions, 3) class separation of location of knot, 4) the mean of the slope growth factor of the second phase, 5) the variance of the slope growth factor of the second phase, and 6) residual variance of the observed variable.

The second objective is to address the issue of model mis-specification. It is important to analyze this issue because applied researchers have to make model selection decisions. Therefore, the current research examines the possibility of extracting spurious latent classes. To achieve this objective a 1-, 2-, and 3-class piecewise LGMMs are fit to the data sets generated under different manipulated conditions using a two class piecewise linear-linear LGMM as a population model. The number of times the correct

model (i.e., 2-class piecewise linear-linear LGMM) is preferred over incorrect models (i.e., 1- and 3- class piecewise linear-linear LGMMs) using the Bayesian Information Criterion (BIC) (Schwarz, 1978) is examined. Detailed information on the simulation design, model estimation, and parameter recovery are provided in Chapter 3.

CHAPTER 3: METHODOLOGY

3.1 Research Design

To develop a two-class piecewise linear-linear LGMM, and to investigate the extent to which the performance of a two-class piecewise linear-linear LGMM is influenced by different population characteristics, a Monte Carlo simulation approach is used. Conditions that are hypothesized to impact the estimation of the knot, along with other model parameters, include: sample size, class mixing proportions, location of the knot, the mean and the variance of the slope growth factor of the second phase, and the residual variance of the observed variable. To evaluate parameter recovery, the proposed model is fit to data generated from a population model with true (known) parameters, and parameter estimates are then compared with their true values. Additionally, the effect of manipulated conditions on the percentage of properly converged replications is analyzed (where a properly converged replication is a replication for which the solution converges with no parameter estimates outside the possible range for that parameter).

Furthermore, to investigate the issue of model mis-specification, 1-, 2-, and 3-class piecewise linear-linear LGMMs are fit to the data sets generated under different manipulated conditions using a 2-class piecewise linear-linear LGMM as a population model. The process of determining the number of latent classes involves the comparison of the BIC indices across the three models. That is, a lower value of the BIC reflects an improvement in fit, hence, a k -class model is selected when the value of the index associated with the k -class model is lower than that of the $k-1$ and $k+1$ class models. The influence of manipulated conditions on the number of times the correct model (i.e., 2-

class piecewise linear-linear LGMM) is preferred over incorrect models (i.e., 1- and 3-class piecewise linear-linear LGMMs) using the BIC criteria is analyzed in Chapter 4.

The BIC is chosen as the model selection index because it is known to pick the correct model most consistently in the framework of finite mixture structure equation models (Jedidi, Jagpal, and DeSarbo, 1997), hence, it has been recommended to be used in the framework of finite mixture modeling (Haughton, 1997; Leroux, 1992). It is based on the loglikelihood (LL) of the postulated model, the number of parameters (p), and sample size (n) as follows:

$$BIC = -2 LL + p \ln (n). \quad (3.1)$$

Another reason for choosing the BIC as a criterion for model selection is that it includes a penalty function for the number of parameters and sample size. That is, BIC gives information about whether a more complicated model fits better than a simpler model over and above their difference in complexity. It is a useful feature because it selects a model that not only fits the data better, but also needs fewer parameters. Other information criteria, such as Akaike's Information Criterion (AIC; Akaike, 1987), also includes a penalty function, but compared to BIC the AIC penalizes models with larger numbers of parameters less, leading to the choice of more mixture components. In other words, the AIC's penalty function is more relaxed as compared to the penalty function of BIC (McLachlan & Peel, 2000). Hence, the AIC tends to overestimate the correct number of mixture components (Celeux & Soromenho, 1996; Soromenho, 1993), whereas the BIC has been reported to perform well (Roeder & Wasserman, 1997). BIC, thus, is the preferred criterion for model comparison because of the advantages that it offers.

The remainder of this chapter describes the data generation procedure, simulation design, including details on the fixed and manipulated conditions, model estimation, and parameter recovery.

3.2 Data Generation Model

A two-class piecewise linear-linear LGMM is used as a population model to generate repeated measures data conforming to 9 equally spaced timepoints (coded 0 to 8) that follow a multivariate normal distribution. The R program (R Development Core Team, 2009) is the statistical package employed to generate the data sets. The choice of two-latent classes for simulation purposes is made so as to keep the scope of the study manageable. It is often seen in both methodological and substantive research of piecewise growth models that the number of timepoints is six or more (see, e.g., Cudeck, 1996; Cudeck & Klebe, 2002; Harring et al., 2006), hence, the choice of 9 timepoints seems to be reasonable. Furthermore, according to research conducted by Lubke and Muthén (2007), additional timepoints do not generally influence model performance or class assignment in the context of linear LGMMs.

Assuming conditional independence, the class-specific population measurement model for the i th individual in the present simulation study is specified as:

$$y_{jk} = \begin{cases} \eta_1 + \eta_2 t_j + \epsilon_j & t_j \leq \gamma_k \\ \eta_1 + \eta_2 \gamma_k + \eta_{4k}(t_j - \gamma_k) + \epsilon_j & t_j > \gamma_k \end{cases}, \text{ for } k = 1, 2, \quad (3.2)$$

where

$$\epsilon_j \sim N(0, \Theta).$$

The specification of Equation 3.2 is similar to that given for piecewise LGC model specified in Equation 2.24. That is, y_{jk} is a set of responses of individual i on a set of

repeated measurement occasions $\mathbf{t}_j = (0,1,2,3,4,5,6,7,8)'$; η_1 and η_2 represent the intercept and slope growth factors for the first phase, respectively; η_{4k} represents the slope growth factor for the second phase; γ_k represents the knot; and ϵ_j is the residual variance. Note that there is no i subscript for any of the model parameters because the data generated will be a balanced and complete data for the individual subjects in the study. Additionally, note that the parameters that do not have k subscript in Equation (3.2) denote population conditions/characteristics that are constrained to be equal across classes for the data generation purposes. The parameters that do have k subscript in Equation (3.2) denote the population conditions that are allowed to vary across classes. Furthermore, it is assumed in this model that the residuals are independent between measurement occasions with constant variance across time. That is,

$$\Theta = \text{diag}(\sigma_\epsilon^2, \sigma_\epsilon^2, \dots, \sigma_\epsilon^2)',$$

where the $\text{diag}(\cdot)$ operator refers to a diagonal matrix, and σ_ϵ^2 refers to the variance of the measurement residuals.

The class-specific structural component of the model for the i th individual is specified as:

$$\begin{pmatrix} \eta_1 \\ \eta_2 \\ \eta_{4k} \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_{4k} \end{pmatrix} + \begin{pmatrix} \zeta_1 \\ \zeta_2 \\ \zeta_4 \end{pmatrix} \quad (3.3)$$

where α is a vector of factor means and ζ is a vector of residuals assumed to be normally distributed with mean zero and covariance matrix Ψ . That is,

$$\zeta \sim N(0, \Psi),$$

where

$$\Psi = \begin{bmatrix} \sigma_{\eta_1}^2 & & \\ 0 & \sigma_{\eta_2}^2 & \\ 0 & 0 & \sigma_{\eta_4}^2 \end{bmatrix},$$

and $\sigma_{\eta_1}^2$, $\sigma_{\eta_2}^2$, and $\sigma_{\eta_4}^2$ are the variances of the intercept growth factor of the first phase, the slope growth factor of the first phase, and the slope growth factor of the second phase, respectively. Note that the parameters that do not have k subscript in Equation (3.3) denote population conditions that are constrained to be equal across classes for the data generation purposes. The parameters that do have k subscript in Equation (3.3) denote the population conditions that are allowed to vary across classes. Furthermore, for the purpose of this study, it is assumed that the intercept and slope growth factors are uncorrelated, thereby simplifying the data generation model. This assumption is consistent with previous studies (see, e.g., Hamilton, 2009).

3.3 Simulation Design

In the data generation process, some population characteristics/conditions (i.e., mean and variance) are held equal across classes, while others are allowed to vary across classes. To elaborate,

Population conditions equal across classes:

1. The mean of the intercept (α_1) and slope growth factors of the first phase (α_2)
2. The variances of the intercept ($\sigma_{\eta_1}^2$) and slope growth factors of the first phase ($\sigma_{\eta_2}^2$), and the variance of the slope growth factor of the second phase ($\sigma_{\eta_4}^2$)
3. The residual variance of the observed variable (σ_{ϵ}^2).

Population conditions not equal across classes:

1. The location of the knot (γ_k)
2. The mean of the slope growth factor of the second phase (α_{4k}).

Note that the decision to keep the population condition for $\sigma_{\eta_1}^2$, $\sigma_{\eta_2}^2$, $\sigma_{\eta_4}^2$, and σ_ϵ^2 equal across both the classes is based on the suggestion made by Muthén (2001) that mixture models with large differences in the factor variances and covariances between classes are particularly sensitive to local maxima.

Furthermore, some conditions are fixed throughout all simulations, while other conditions are manipulated. The fixed and manipulated conditions are described in the following sections.

3.2.1 Fixed conditions

The focal point of this research is to estimate the location of the knot across classes in different manipulated conditions, thus the conditions that were not directly relevant to the study of the knots were fixed across all simulations. The fixed conditions are:

1. Population mean of the intercept growth factor of the first phase (α_1)
2. Population variance of the intercept growth factor of the first phase ($\sigma_{\eta_1}^2$)
3. Population mean of the slope growth factor of the first phase (α_2)
4. Population variance of the slope growth factor of the first phase ($\sigma_{\eta_2}^2$)
5. The factor covariances are fixed to zero.

The population mean trajectory within each class is parameterized so that, on average, scores will not increase over time until $t_{ji} > \gamma_k$ (i.e., $\alpha_2 = 0$). Additionally, the population values of α_1 , $\sigma_{\eta_1}^2$, and $\sigma_{\eta_2}^2$ are chosen so that they are similar to what is

commonly found in previous simulation studies in the area of LGMM. The population values for the fixed conditions are provided in Table 1.

Table 1.

Population values for fixed conditions.

Conditions	Population Values
Growth Factor Means	
• Intercept 1 (α_1)	2.0
• Slope 1 (α_2)	0.0
Growth Factor Variances	
• Intercept 1 ($\sigma_{\eta_1}^2$)	1.0
• Slope 1 ($\sigma_{\eta_2}^2$)	0.2
Growth Factor Covariances	
• Intercepts and slopes	0.0

The population value of $\alpha_1 = 2.0$ is similar to the value used by Nylund, Asparouhov, and Muthén (2007). The population values of $\sigma_{\eta_1}^2$ and $\sigma_{\eta_2}^2$ are selected so that they are in the ratio of 5:1 (i.e., $\frac{\sigma_{\eta_1}^2}{\sigma_{\eta_2}^2} = \frac{1}{0.20} = \frac{5}{1}$). This ratio is consistent with the ratio used in previous LGMM simulation studies (see, e.g., Bauer & Curran, 2003; Hamilton, 2009).

3.2.1 Manipulated conditions

The manipulated conditions are summarized in Table 2 before being described more fully in the following sections.

Table 2.

Manipulated conditions.

Conditions	Number of Levels
• Sample Size	3
• Class Mixing Proportions (φ_k)	2
• Class Separation of Location of Knot (γ_k)	5
• Growth Factor Mean of Slope of Second Phase (α_{4k})	3
• Growth Factor Variance of Slope of Second Phase ($\sigma_{\eta_4}^2$)	2
• Residual Variance of Observed Variables (σ_{ϵ}^2)	2

1. Sample size, n

The three levels of sample size chosen are: 400, 700, and 1000. These three levels are chosen based on the results obtained from a pilot study conducted to determine the three best levels of sample size out of a total of five different sample sizes (i.e., 200, 400, 700, 1000, and 2000). The detailed discussion on the results obtained from the pilot work is presented in a subsequent section.

2. Class mixing proportion (φ_k)

The two levels of mixing proportion chosen are: (50/50 and 75/25) based on Nylund et al. (2007).

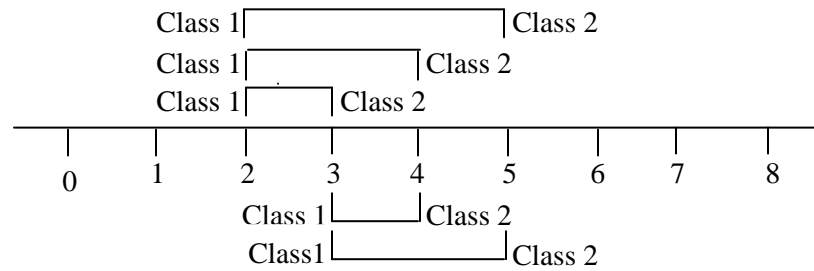
3. Class separation of location of knot (γ_k)

There are 9 equally spaced timepoints in the generated data sets. The range in which the population values of the knot for the two classes are chosen is between timepoint 2 and timepoint 6. This is so because before timepoint 2 and after timepoint 6 there is too little information available to estimate the mean and the variance of the slope of the first phase and of the slope of the second phase,

respectively. The levels of the location of knot condition are determined in the following way:

		Class 1 Knot Location		
		2	3	4
Class 2 Knot Location	+1	2, 3	3, 4	4, 5
	+2	2, 4	3, 5	4, 6
	+3	2, 5	3, 6	4, 7

Out of these nine cells, only those cells are chosen as levels of the class separation condition that do not create mirror images. Hence, based on this criterion five cells are chosen as levels of this condition; that is, (2, 3); (2, 4); (2, 5); (3, 4); and (3, 5). The conditions can also be illustrated in the following way:



4. The population mean of the slope growth factor of the second phase (α_{4k})

The three levels of this condition are:

- a) low α_{4k} in both classes 1 and 2
- b) low α_{41} in class 1, and high α_{42} in class 2
- c) high α_{4k} in both classes 1 and 2.

The population mean values of α_{4k} corresponding to the above stated three conditions, respectively, are: (0.25, 0.25); (0.25, 1.25); and (1.25, 1.25). The rationale behind the choice of this condition is that the degree of bend between the

two slopes (i.e., η_{2k} and η_{4k}) could affect the estimation of location of knot, and hence the ability to distinguish between classes having different knots.

5. The variance of the slope growth factor of the second phase ($\sigma_{\eta_4}^2$)

The two levels of this condition are: low variance ($\sigma_{\eta_4}^2 = 0.20$) and high variance ($\sigma_{\eta_4}^2 = 1.0$) relative to the variance of the slope growth factor of the first phase (i.e., $\sigma_{\eta_2}^2 = 0.20$). Note that it has been stated above that the variances of the slope growth factor of the second phase are held equal across classes in the data generation process.

6. Residual variance of the observed variable (σ_{ϵ}^2)

The two levels of this condition are: low variance (i.e., $\sigma_{\epsilon}^2 = 1.0$) and high variance (i.e., $\sigma_{\epsilon}^2 = 5.0$). This condition is selected based on the common knowledge that the amount of residual variance in the observed variable can affect the fitting of the model, thereby affecting the estimation of the model parameters, including location of the knot. In other words, when the amount of residual variance in observed variable is small, it should be relatively easy to fit the function, thereby making it possible to estimate the location of the knot in the fitted function. But when the amount of variance is large, it may be difficult to fit the function, thereby making it relatively difficult to estimate the location of the knot. Hence, this condition is relevant in the context of estimation of knot location. Furthermore, the population values for the low and high condition is selected keeping in mind the range of intraclass correlation commonly found in practice (i.e., ≤ 0.06). The intraclass correlation coefficient represents the

proportion of variance in an outcome variable explained by between subject variability, that is,

$$\rho_{\text{ICC}} = \left[\frac{\textit{Between subject variability}}{\textit{Between subject variability} + \textit{Within subject variability}} \right]$$

which translates to,

$$\rho_{\text{ICC}} = \left[\frac{\textit{Variance of intercept factor of the first phase}}{\textit{Variance of intercept factor of the first phase} + \textit{Residual variance}} \right].$$

Also note, it has been stated earlier that the residual variance of the observed variables is held equal across classes in the data generation process.

The combination of manipulated conditions ($3 \times 2 \times 5 \times 3 \times 2 \times 2$) results in a Monte Carlo simulation with 360 cells. For each of the cells, 100 replications are generated to assess the results obtained. Table 3 reports the population values for fixed and manipulated conditions that are used to generate two-class piecewise linear-linear LGMMs.

Table 3.

Fixed and manipulated parameter values.

		Class 1	Class 2
Fixed Conditions	Growth Factor Means of Intercept and Slope of First Phase		
	• Intercept 1 (α_1)	2.0	2.0
	• Slope 1 (α_2)	0.0	0.0
	Growth Factor Variances of Intercept and Slope of First Phase		
	• Intercept 1 ($\sigma_{\eta_1}^2$)	1.0	1.0
	• Slope 1 ($\sigma_{\eta_2}^2$)	0.2	0.2
Manipulated Conditions	Growth Factor Covariances		
	• Intercepts and slopes	0.0	0.0
	Class Separation of Location of Knot (γ_k)		
	• Level 1	2	3
	• Level 2	2	4
	• Level 3	2	5
	• Level 4	3	4
	• Level 5	3	5
	Growth Factor Mean of Slope of Second Phase (α_{4k})		
	• Level 1: Low-Low	0.25	0.25
	• Level 2: Low-High	0.25	1.25
	• Level 3: High-High	1.25	1.25
	Growth Factor Variance of Slope of Second Phase ($\sigma_{\eta_4}^2$)		
	• Level 1: Low-Low	0.2	0.2
	• Level 2: High-High	1.0	1.0
	Residual Variance of Observed Variables (σ_ϵ^2)		
	• Level 1: Low-Low	1.0	1.0
	• Level 2: High-High	5.0	5.0

3.4 Pilot Analysis

3.4.1 Selection of Levels of Sample Size Condition

There was no previous study on piecewise linear-linear LGMM, hence, as a first step a small pilot simulation was conducted to determine the three levels of sample size on the basis of convergence criterion, and whether there were important differences in results from different sample sizes, say, 1000 to 2000. The critical range of sample size is: 200, 400, 700, 1000, and 2000.

Data were generated using a two-class piecewise linear-linear LGMM with 9 equally spaced timepoints (coded 0 to 8) for each of the sample size in the critical range under six manipulated conditions. The manipulated conditions included, only three levels of class separation of location of knot [i.e., (2, 3), (3, 5), and (2, 5)], and only two levels of population mean of the slope growth factor of the second phase (i.e., low α_{4k} in both classes 1 and 2, and low α_{41} in class 1, and high α_{42} in class 2). The values used for low α_{4k} in both classes 1 and 2, and low α_{41} in class 1, and high α_{42} in class 2 are those described in Table 3. The remaining conditions were fixed: class mixing proportion (75/25), growth factor variance of slope of second phase ($\sigma_{\eta_4}^2 = 1.0$), and residual variance of observed variables ($\sigma_{\epsilon}^2 = 5.0$). The fixed parameter values that were used are those described in Table 1. The combination of manipulated conditions and the five levels of sample size resulted in a Monte Carlo simulation with 30 cells. For each of the cells, 100 replications were generated to assess the results obtained. The models were estimated using *Mplus* 6.1 (Muthén & Muthén, 1998-2010) where population values were provided as start values for the parameters to be estimated. The default estimator for mixture analysis using *Mplus* is ML via the EM algorithm.

An examination of Tables C1 through C6 in Appendix C yields the following key observations:

- For sample size $n = 200$, the proportion of converged replications in each cell was lowest when compared with other sample sizes in the range.
- There were no important differences in outcome measures (i.e., average parameter bias and standard deviation around average parameter bias) from sample sizes $n = 1000$ and $n = 2000$.
- Overall, the model seemed to perform well when sample size $n = 400, 700, 1000$, and 2000 in terms of the measured outcomes and convergence rate.

Thus, the three levels of sample size selected were $n = 400, 700$, and 1000 .

3.4.2 Asymptotic Behavior of Model Parameters for a 2-Class Piecewise Linear-Linear LGMM

A small pilot study was conducted to analyze the asymptotic behavior of model parameters for a two-class piecewise linear-linear LGMM using ML estimation via the EM algorithm. Given a very large sample size, it is expected that ML estimation via the EM algorithm will produce model parameter estimates that are close to the known (true) population values. Data for this pilot study were generated using a two class piecewise linear-linear LGMM with 9 equally spaced timepoints for sample size, $n = 100,000$. The population values for the data generation of the piecewise linear-linear LGMM are provided in Table C7 in Appendix C. The model was estimated using *Mplus* 6.1 where population values were provided as start values for the parameters to be estimated.

As seen in Table C8 in Appendix C, the estimated original model parameters of a two-class piecewise linear-linear LGMM are almost same as the known (true) population values. Thus, it can be concluded that ML estimation via the EM algorithm does a reasonable of estimating model parameters.

3.5 Model Estimation

The parameters of two-class piecewise linear-linear LGMM are estimated using *Mplus* 6.1. The default estimator for mixture analyses using *Mplus* is ML via the EM algorithm. According to the research conducted by Lubke and Muthén (2007), the complexity of the model with respect to the factor structure, or the number of observed variables within class, do not influence model performance. However, when estimating mixture models, in general, using ML estimation via EM algorithm, failure to converge to a stable solution within a given number of iterations or converging to a local maximum of the likelihood are common problems (Bauer & Curran, 2003; Muthén, 2001).

Muthén and Muthén (2000) suggested that researchers provide starting values of the parameters to be estimated that reflects their beliefs about the population as it helps the modeling algorithm to converge. As this study was not intended to focus on convergence issues, the population values of the parameters are used as the starting values in *Mplus* 6.1. The decision of choosing population values as starting values is consistent with previous studies (see, e.g., Hamilton, 2009; Paxton, Curran, Bollen, Kirby, & Chen, 2001). In addition, the number of default starting values in *Mplus* 6.1 (i.e., ten sets of random starting values are used with two iterations for each set) are

increased to fifty sets of random starting values with five iterations for each set in order to investigate local solutions more thoroughly.

While estimating piecewise linear-linear LGMMs under different manipulated conditions in *Mplus* 6.1, only the residual variance of observed variable is constrained to be equal across classes. Rest all the other model parameters are allowed to be freely estimated across classes. That is, the mean and the variance of the intercept growth factor of the first phase, the mean and the variance of the slope growth factor of the first phase, the mean and the variance of the slope growth factor of the second phase, the growth factor covariances (intercepts and slopes) and the location of the knot are estimated for each class. Once the models have been estimated in *Mplus* 6.1, the parameter estimates of interest are imported into the R program (R Development Core Team, 2009) for further analyses as described in the next section.

3.6 Outcome Measures

Upon convergence, the estimated parameters of the reparameterized models are transformed back to the original parameters of the two-class piecewise linear-linear LGMMs using the procedure shown in Appendix B. The transformation is carried out in the R program. To evaluate the performance of two-class piecewise linear-linear LGMMs under different manipulated conditions, the following outcome measures are used: parameter bias, and variability index for parameter bias. Parameter bias is defined as the difference between the estimated parameter value and the corresponding population true value, that is,

$$\text{Bias} = [\hat{\theta}_k - \theta_{\text{True}}].$$

Note that Bias is computed for only successful replications (i.e., a replication for which the solution converged with no parameter estimates outside the possible range for the parameter) in each cell. Additionally, variability index for parameter bias corresponding to each of the estimated parameters for each replication in each cell is computed as:

$$\text{Variability Index} = \left[\hat{\theta}_k - \bar{\theta} \right]^2.$$

To evaluate the accuracy of the parameter estimates, percent bias for each of the estimated model parameters in each cell is computed. A percent bias is defined as:

$$\text{Percent Bias} = \left[\frac{\text{Median bias for the } p^{\text{th}} \text{ parameter}}{\text{Population value for the } p^{\text{th}} \text{ parameter}} \right] \times 100.$$

Note that median bias was used in the above equation instead of mean bias because it is more resistant to outliers. Positive values for percent bias occur for estimates that are above the population value by the percent magnitude listed, whereas negative values for percent bias occur for estimates that are below the population value by the percent magnitude listed. Parameter estimates that contained 10% bias or more in either direction are considered definitely biased (Gagné, 2004) and are reported in Chapter 4.

Furthermore, to quantify bias as a function of the manipulated conditions, analysis of variance (ANOVA) with a 6-way [3 (sample size) \times 2 (class mixing proportion) \times 5 (class separation of location of knot) \times 3 (growth factor mean of slope of second phase) \times 2 (growth factor variance of slope of second phase) \times 2 (residual variance in observed variables)] is performed. Partial eta squared, η^2 , corresponding to each manipulated factor and the interaction terms, are reported for practical significance. Partial η^2 for a manipulated factor is defined as the proportion of total variation attributable to the factor,

partialling out (excluding) other factors from the total nonerror variation (Pierce, Block, & Aguinis, 2004, p. 918).

The criterion used for characterizing η^2 values is the same as that for characterizing η^2 values. Using the same criterion is legitimate because with large samples, the distinction between η^2 and η^2 tends to be small as η^2 involves division by total sample size (Sapp, 2006). Partial η^2 involves division by sample size minus number of groups (Sapp, 2006). Since the sample size for the ANOVA analyses in the context of this study is large, it is reasonable to use the criterion for characterizing η^2 values as the criterion for characterizing η^2 values. η^2 values are characterized as small, medium, or large, where .01 constitutes small, .06 medium, and 0.14 large (see Cohen, 1988, p. 283). Note that the main effects of the manipulated factors and the interaction terms are reported and interpreted only when both statistical significance ($p < .05$) and practical significance ($\eta^2 \geq 0.06$) are achieved.

The following chapter presents results related to the first and the second research objectives of the current study. Results are presented regarding the influence of manipulated conditions on the properly converged replications for the two-class piecewise linear-linear LGMMs, and on the model selection rate. Results are also presented related to the accuracy of estimated parameters from the two-class piecewise linear-linear LGMM with data generated under known study conditions.

CHAPTER 4 – RESULTS

In this chapter, results of the simulation study are organized and presented in order to address the first and the second research objectives of this study. Results related to the influence of manipulated conditions on the properly converged replications for the two-class piecewise linear-linear LGMM are presented in section 4.1; results related to the parameter bias and the variability index of parameter bias from estimating parameters from the two-class piecewise linear-linear LGMM with data generated under known study conditions are presented in section 4.2; results from model mis-specification are presented in section 4.3; and a summary of main findings is presented in section 4.4.

4.1 Model Convergence Rate

The rate of converged replications for two-class piecewise linear-linear LGMMs across all the cells (total number of cells were 360) was found to be between 58% and 99%, where 166 cells out of 360 had a 90% or higher rate of convergence to the global solution, 100 cells out of 360 had a 70% to 89% rate of convergence to the global solution, and 94 cells out of 360 had a 69% or lower rate of convergence to the global solution. The proportion of properly converged replications for two-class piecewise linear-linear LGMMs are shown in Table 4. Note that the Table 4 has been arranged from ascending to descending order using proportions of properly converged replications as the sorting variable.

Table 4.

Proportion of properly converged replications for two-class piecewise linear-linear LGMMs.

Cell	Prop.	Cell	Prop.	Cell	Prop.	Cell	Prop.	Cell	Prop.	Cell	Prop.
1	.58	39	.63	77	.67	115	.75	153	.82	191	.89
2	.58	40	.63	78	.67	116	.75	154	.82	192	.89
3	.58	41	.63	79	.67	117	.75	155	.82	193	.89
4	.59	42	.63	80	.67	118	.75	156	.82	194	.89
5	.59	43	.63	81	.67	119	.76	157	.83	195	.90
6	.59	44	.63	82	.67	120	.76	158	.83	196	.90
7	.59	45	.63	83	.67	121	.76	159	.83	197	.90
8	.59	46	.63	84	.67	122	.76	160	.83	198	.90
9	.60	47	.63	85	.67	123	.76	161	.84	199	.90
10	.60	48	.63	86	.68	124	.76	162	.84	200	.90
11	.60	49	.63	87	.68	125	.77	163	.84	201	.90
12	.60	50	.64	88	.68	126	.77	164	.84	202	.90
13	.60	51	.64	89	.68	127	.77	165	.84	203	.90
14	.60	52	.64	90	.68	128	.77	166	.84	204	.90
15	.60	53	.64	91	.69	129	.77	167	.84	205	.91
16	.60	54	.64	92	.69	130	.78	168	.85	206	.91
17	.61	55	.64	93	.69	131	.78	169	.85	207	.91
18	.61	56	.64	94	.69	132	.78	170	.86	208	.91
19	.61	57	.64	95	.70	133	.78	171	.86	209	.91
20	.61	58	.64	96	.70	134	.78	172	.86	210	.91
21	.61	59	.64	97	.70	135	.79	173	.86	211	.91
22	.61	60	.64	98	.70	136	.79	174	.87	212	.91
23	.61	61	.65	99	.71	137	.79	175	.87	213	.91
24	.62	62	.65	100	.71	138	.79	176	.87	214	.91
25	.62	63	.65	101	.71	139	.79	177	.87	215	.91
26	.62	64	.65	102	.72	140	.79	178	.87	216	.91
27	.62	65	.65	103	.72	141	.79	179	.88	217	.91
28	.62	66	.65	104	.72	142	.79	180	.88	218	.91
29	.62	67	.65	105	.73	143	.79	181	.88	219	.91
30	.62	68	.65	106	.73	144	.79	182	.88	220	.91
31	.62	69	.65	107	.73	145	.80	183	.88	221	.91
32	.63	70	.65	108	.73	146	.80	184	.88	222	.92
33	.63	71	.66	109	.73	147	.80	185	.89	223	.92
34	.63	72	.66	110	.73	148	.81	186	.89	224	.92
35	.63	73	.66	111	.73	149	.81	187	.89	225	.92
36	.63	74	.66	112	.73	150	.82	188	.89	226	.92
37	.63	75	.66	113	.74	151	.82	189	.89	227	.92
38	.63	76	.66	114	.74	152	.82	190	.89	228	.92

Table 4 (contd.)

Proportion of properly converged replications for two-class piecewise linear-linear LGMMs.

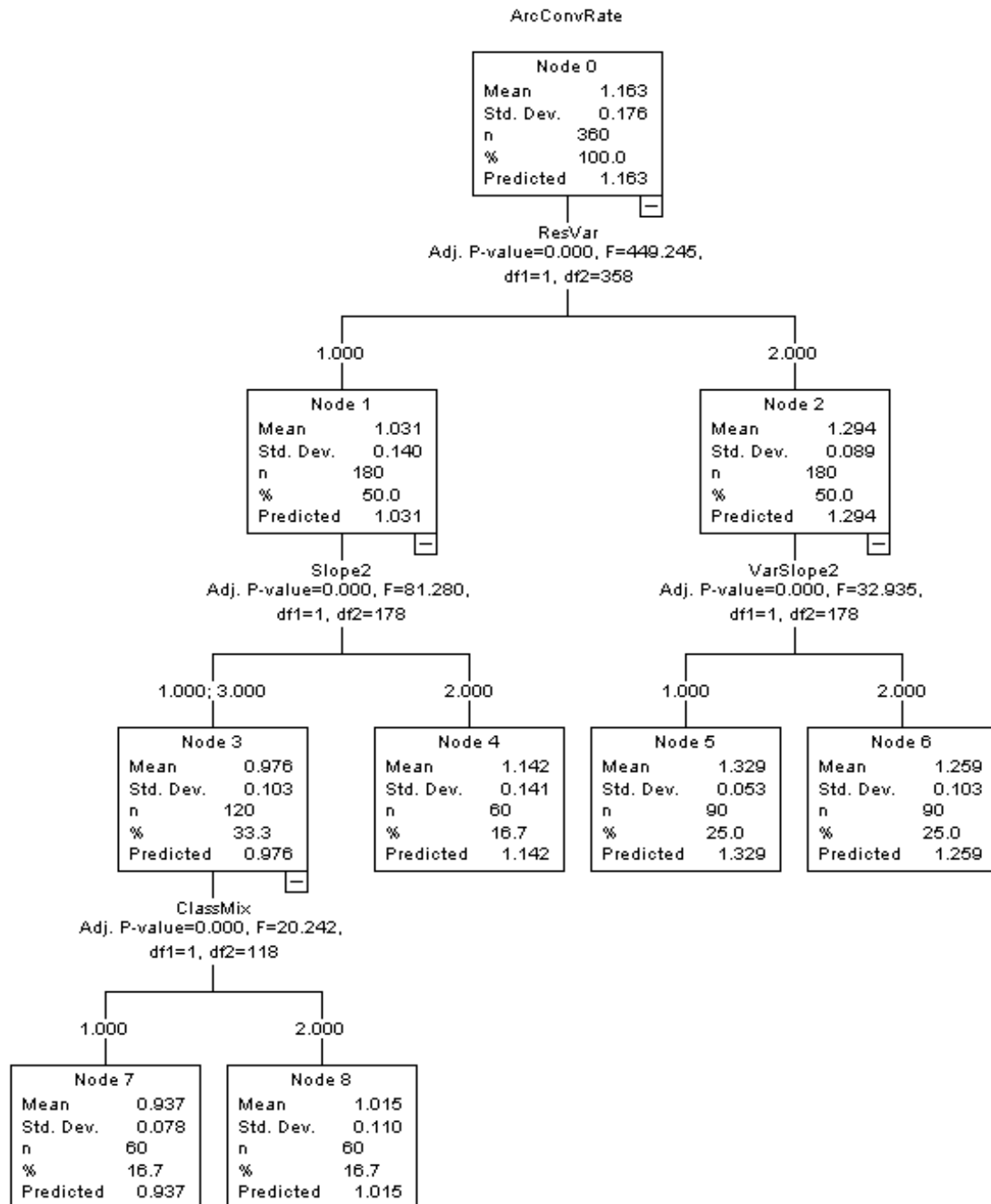
Cell	Prop.	Cell	Prop.	Cell	Prop.	Cell	Prop.
229	.92	267	.94	305	.95	343	.97
230	.92	268	.94	306	.95	344	.97
231	.92	269	.94	307	.95	345	.97
232	.92	270	.94	308	.95	346	.97
233	.92	271	.94	309	.95	347	.97
234	.92	272	.94	310	.95	348	.97
235	.92	273	.94	311	.96	349	.97
236	.92	274	.94	312	.96	350	.97
237	.92	275	.94	313	.96	351	.97
238	.92	276	.94	314	.96	352	.97
239	.92	277	.94	315	.96	353	.97
240	.92	278	.94	316	.96	354	.97
241	.92	279	.94	317	.96	355	.98
242	.92	280	.94	318	.96	356	.98
243	.92	281	.94	319	.96	357	.98
244	.93	282	.94	320	.96	358	.98
245	.93	283	.94	321	.96	359	.98
246	.93	284	.94	322	.96	360	.99
247	.93	285	.95	323	.96		
248	.93	286	.95	324	.96		
249	.93	287	.95	325	.96		
250	.93	288	.95	326	.96		
251	.93	289	.95	327	.96		
252	.93	290	.95	328	.96		
253	.93	291	.95	329	.96		
254	.93	292	.95	330	.96		
255	.93	293	.95	331	.96		
256	.93	294	.95	332	.96		
257	.93	295	.95	333	.96		
258	.93	296	.95	334	.96		
259	.93	297	.95	335	.96		
260	.93	298	.95	336	.96		
261	.93	299	.95	337	.96		
262	.93	300	.95	338	.96		
263	.93	301	.95	339	.96		
264	.94	302	.95	340	.96		
265	.94	303	.95	341	.96		
266	.94	304	.95	342	.97		

The percentage of properly converged replications was examined to see if it appeared to be a function of the manipulated conditions. The classification tree in Figure 8 graphically depicts the effect of the manipulated conditions on the model convergence rate before being fully described later. The method used for creating classification tree was Chi-squared Automatic Interaction Detection (CHAID). CHAID is an exploratory data analysis method used to study the relations between a dependent variable (i.e., model convergence rate) and independent variables (i.e., manipulated conditions). At each step in the tree creation, CHAID chooses the independent variable that has the strongest statistically significant relation with the dependent variable. The procedure automatically excludes any variables that do not make a significant contribution to the final model. Note that before creating the classification tree, the dependent variable (model convergence rate expressed in proportions) was transformed using an arcsine transformation (units expressed in radians) (Sokal & Rohlf, 1995). This was done because data that are in percents or proportions are generally not normally distributed. To make percent/proportion data closer to normal, an arcsine transformation of data is useful, that is

$$p' = \arcsin\sqrt{p}$$

where p denotes data that are proportions.

Figure 8: Classification Tree: The effect of manipulated conditions on the percentage of properly converged replications.



As observed in Figure 8, the most influential predictor of model convergence rate was the manipulated condition of residual variance in observed variables (σ_{ϵ}^2). In other words, residual variance in observed variables had the strongest statistically significant relation with model convergence rate. There were different levels of model convergence rate for the two nodes, Node 1 (low residual variance, $\sigma_{\epsilon}^2 = 1.0$) and Node 2 (high residual variance, $\sigma_{\epsilon}^2 = 5.0$), formed on the basis of different conditions of residual variance. Node 2 seemed to favor a slightly better model convergence rate as compared to Node 1 (i.e., $\bar{X}_{Node\ 2} = 1.294$; $\bar{X}_{Node\ 1} = 1.031$). This implies that the model convergence rates for the two-class piecewise linear-linear LGMMs were fairly high when the value of residual variance in observed variable was large.

For low residual variance category (Node 1), the next best predictor was the manipulated condition of growth factor mean of the slope of the second phase (α_{4k}). Node 1 was statistically significantly split into two nodes, Node 3 (includes level 1 - low-low: 0.25, 0.25 and level 3 - high-high: 1.25, 1.25) and Node 4 (level 2 - low-high: 0.25, 1.25), on the basis of different conditions of growth factor mean of the slope of the second phase. Node 4 seemed to favor a relatively higher model convergence rate as compared to Node 3 (i.e., $\bar{X}_{Node\ 4} = 1.142$; $\bar{X}_{Node\ 3} = 0.976$). This implies that the cells that combined low value of residual variance and a low-high condition for the growth factor mean of the slope of the second phase favored a relatively higher model convergence rate as compared to the cells that combined low value of residual variance with the low-low or high-high conditions of the growth factor mean of the slope of the second phase.

Furthermore, for Node 3, the next best predictor was the manipulated condition of class mixing proportion condition (φ_k). Node 3 was statistically significantly split into two nodes, Node 7 (level 1 - 50/50) and Node 8 (level 2 - 75/25), on the basis of different conditions of class mixing proportion. Node 8 seemed to favor a relatively higher model convergence rate as compared to Node 7 (i.e., $\bar{X}_{Node\ 8} = 1.015$; $\bar{X}_{Node\ 7} = 0.937$). This implies that the cells that combined low residual variance, the low-low or high-high conditions of the growth factor mean of the slope of the second phase, and 75/25 class mixing proportion favored a relatively higher model convergence rate as compared to the cells that combined low residual variance, the low-low or high-high conditions of the growth factor mean of the slope of the second phase, and 50/50 class mixing proportion.

For high residual variance category (Node 2), the next best predictor was the manipulated condition of growth factor variance of the slope of the second phase ($\sigma_{\eta_4}^2$). Node 2 was statistically significantly split into two nodes, Node 5 (low slope2 variance, $\sigma_{\eta_4}^2 = 0.20$) and Node 6 (high slope2 variance, $\sigma_{\eta_4}^2 = 1.0$), on the basis of different conditions of growth factor variance of the slope of the second phase. Node 5 seemed to favor a relatively higher model convergence rate as compared to Node 6 (i.e., $\bar{X}_{Node\ 5} = 1.329$; $\bar{X}_{Node\ 6} = 1.259$). This implies that the cells that had a combination of high value of residual variance and low slope2 variance favored a relatively higher model convergence rate as compared to the cells that had a combination of high value of residual variance and high slope2 variance. Overall, the cells with high value of residual variance and low slope2 variance had the best model convergence rate.

The result at the end of this tree building process is that we have a series of Nodes defined by the manipulated conditions: residual variance in observed variable (σ_{ϵ}^2),

growth factor mean of the slope of the second phase (α_{4k}), growth factor variance of the slope of the second phase ($\sigma_{\eta_4}^2$), and class mixing proportion (φ_k) that are maximally different from one another on the model convergence rate. The manipulated conditions that did not make a significant contribution to the final model were sample size (n) and class separation of location of knot (γ_k).

4.2 Parameter Bias

Parameter estimates within 10% of the population value were considered acceptable. Using this criterion, none of the parameter estimates for growth factor mean of the intercept of the first phase, and the residual variance in observed variable for class 1 and 2, respectively, across all 360 cells were considered unacceptable. The percentage of cells that had unacceptable values of the parameter estimates for growth factor mean of the slope of the second phase for class 1 and 2 were 32.50% and 25%, respectively. Additionally, nearly 30% of the cells had unacceptable values of the parameter estimates for location of the knot for class 1 and 2.

Furthermore, the parameter estimates for variances of the growth factors (intercept of the first phase, slope of the first phase, and slope of the second phase) were generally poor for both the classes. The percentage of cells that had unacceptable values of parameter estimates for variances of growth factors ranged between 61.11% and 100% for class 1 and 2, respectively. The percentage of cells with 10% or more bias in either direction is summarized in Table 5.

Table 5.

Percentage of cells with 10% or more bias in either direction.

% of Cells with %Bias ≥ 10	
Parameter Bias (Class 1)	Bias_ α_{11}
	0.00
	Bias_ α_{21}
	N/A
	Bias_ α_{41}
	32.50
	Bias_ γ_1
	29.44
Parameter Bias (Class 2)	Bias_ $\sigma_{\eta_{11}}^2$
	100.00
	Bias_ $\sigma_{\eta_{21}}^2$
	61.11
	Bias_ $\sigma_{\eta_{41}}^2$
	61.94
	Bias_ $\sigma_{\epsilon_1}^2$
	0.00
Parameter Bias (Class 2)	Bias_ α_{12}
	0.00
	Bias_ α_{22}
	N/A
	Bias_ α_{42}
	25.00
	Bias_ γ_2
	28.33
Parameter Bias (Class 2)	Bias_ $\sigma_{\eta_{12}}^2$
	100.00
	Bias_ $\sigma_{\eta_{22}}^2$
	61.94
	Bias_ $\sigma_{\eta_{42}}^2$
	65.83
	Bias_ $\sigma_{\epsilon_2}^2$
	0.00

Note: that the population value of the growth factor mean of slope of the first phase is 0 for both the classes across all the cells. Thus, it is not possible to compute the %bias for this particular parameter.

To quantify bias as a function of the manipulated conditions a 6-way analysis of variance (ANOVA) [3 (sample size) \times 2 (class mixing proportion) \times 5 (class separation

of location of knot) \times 3 (growth factor mean of slope of second phase) \times 2 (growth factor variance of slope of second phase) \times 2 (residual variance in observed variable)] was performed on the outcome measures: parameter bias and variability index of parameter bias. The main effects of the manipulated conditions and the interaction terms were reported and interpreted only when both statistical significance ($p < .05$) and practical significance ($\eta^2 \geq 0.06$) were achieved.

In all the ANOVA tables there were no main effects or interaction terms that satisfied both the statistical and the practical significance criteria. Thus, it can be concluded that the parameter bias and the variability index of parameter bias are not systematically related to any of the manipulated conditions in the study. A summary of results obtained from the ANOVA tables are presented in Tables D1 and D2 in Appendix D.

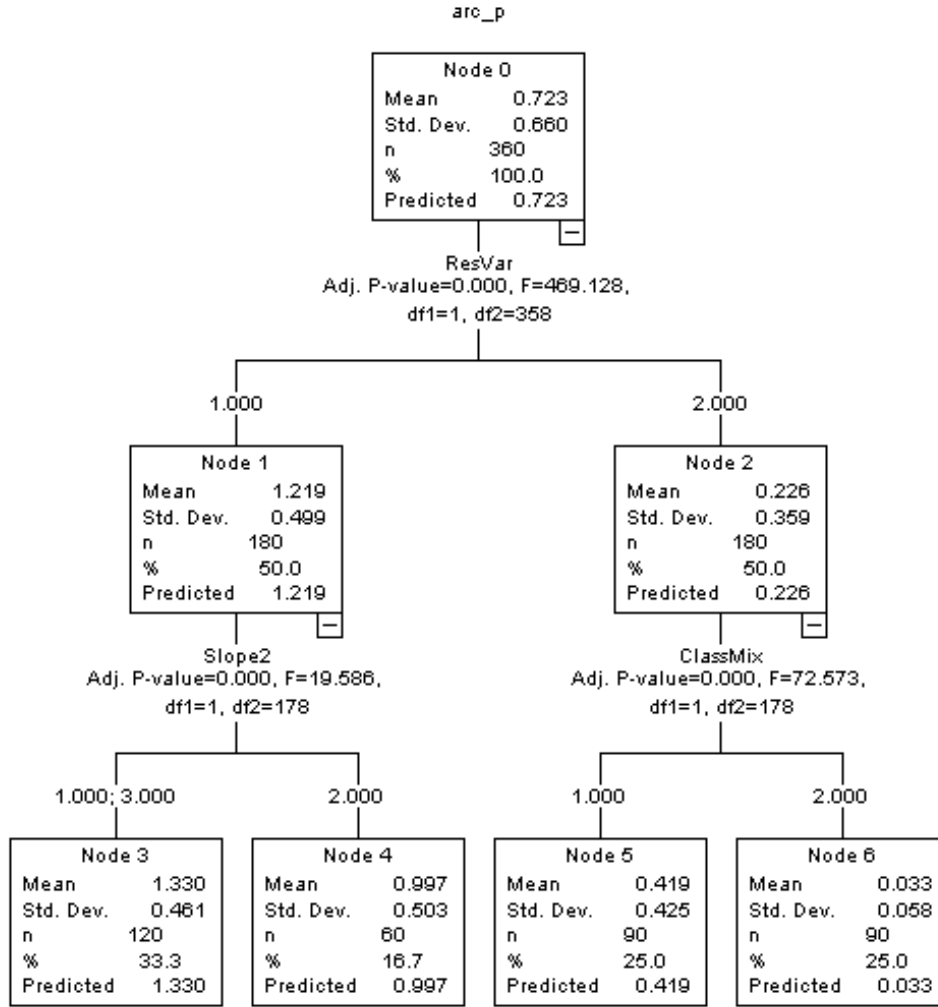
4.3. Model Mis-specification

The 3-class piecewise linear-linear LGMM failed to converge for all the replications across all the cells, so it was not possible to obtain the BIC indices from these replications. This is not too surprising given that, unlike the 2-class piecewise linear-linear LGMM, starting values for this model could not be provided in *Mplus* 6.1. Furthermore, the review of the literature indicates that typically over-extracted latent class models have serious convergence problems (Nylund et al., 2007; Tofighi & Enders, 2008). Thus, the failed replications of 3-class piecewise linear-linear LGMM were simply discarded and the analyses were based on the replications that produced a converged solution for the 1- and 2-class piecewise linear-linear LGMM.

4.3.1. Model selection rate

Model selection rates across all the cells (total number of cells were 360) were found to be between .00 and 1.00, where 139 cells out of 360 had a 80% or higher rate of model selection, 50 cells out of 360 had a 21% to 79% rate, and 171 cells out of 360 had a 20% or lower rate of model selection. The number of times the correct model (i.e., 2-class piecewise linear-linear LGMM) was preferred over the incorrect model (i.e., 1-class piecewise linear-linear LGMM) was examined to see if it appeared to be a function of the manipulated conditions. The classification tree in Figure 9 graphically depicts the effect of the manipulated conditions on the model selection rate before being fully described later. The method used for creating the classification tree was Chi-squared Automatic Interaction Detection (CHAID). Note that before creating the classification tree, the dependent variable (model selection rate expressed in proportions) was transformed using an arcsine transformation (units expressed in radians). This transformation was done so as to make the distribution of model selection rate data more normal.

Figure 9: Classification Tree: The effect of manipulated conditions on the selection rate.



As seen in Figure 9, the most influential predictor of model convergence rate was the manipulated condition of residual variance in observed variables (σ_{ϵ}^2). In other words, residual variance in observed variables had the strongest statistically significant relation with model selection rate. There were different levels of model selection rate for Node 1 (low residual variance, $\sigma_{\epsilon}^2 = 1.0$) and Node 2 (high residual variance, $\sigma_{\epsilon}^2 = 5.0$)

formed on the basis of different conditions of residual variance in observed variable.

Node 1 seemed to favor a much higher model selection rate as compared to Node 2 (i.e., $\bar{X}_{Node\ 1} = 1.219$; $\bar{X}_{Node\ 2} = 0.226$). This implies that the model selection rates for the two-class piecewise linear-linear LGMMs were high when the value of residual variance in observed variable was small.

For low residual variance category (Node 1), the next best predictor was the manipulated condition of growth factor mean of the slope of the second phase (α_{4k}). Node 1 was statistically significantly split into two nodes, Node 3 (includes level 1 - low-low: 0.25, 0.25 and level 3 - high-high: 1.25, 1.25) and Node 4 (level 2 - low-high: 0.25, 1.25), on the basis of the growth factor mean of the slope of the second phase. Node 3 seemed to favor a higher model selection rate as compared to Node 4 (i.e., $\bar{X}_{Node\ 3} = 1.330$; $\bar{X}_{Node\ 4} = 0.997$). This, basically, implies that the cells that combined the low-low or high-high conditions of the growth factor mean of the slope of the second phase and low value of residual variance favored a higher model selection rate as compared to the cells that had a combination of low value of residual variance and the low-high condition of the growth factor mean of the slope of the second phase. Overall, the cells with low value of residual variance and the low-low or high-high conditions of the growth factor mean of the slope of the second phase had the best model selection rate.

For high residual variance category (Node 2), the next best predictor was the manipulated condition of class mixing proportion (φ_k). Node 2 was statistically significantly split into two nodes, Node 5 (level 1 - 50/50) and Node 6 (level 2 - 75/25), on the basis of class mixing proportion. Node 5 seemed to favor a higher model selection rate as compared to Node 6 (i.e., $\bar{X}_{Node\ 5} = 0.419$; $\bar{X}_{Node\ 6} = 0.033$). This implies that

the cells that had high value of residual variance and 50/50 class mixing proportion favored a relatively higher model selection rate as compared to the cells that had a combination of high residual variance and 75/25 class mixing proportion.

The result at the end of this tree building process is that we have a series Nodes defined by the manipulated conditions of residual variance of observed variable (σ_{ϵ}^2), growth factor mean of the slope of the second phase (α_{4k}), and class mixing proportion (ϕ_k) that are maximally different from one another on the model selection rate. The manipulated conditions that didn't make a significant contribution to the final model were sample size (n), class separation of location of knot (γ_k), and growth factor variance of the slope of the second phase ($\sigma_{\eta_4}^2$).

4.3.2. Parameter bias and variability index of parameter bias for filtered replications

As a post hoc analysis, a 6-way analysis of variance (ANOVA) [3 (sample size) \times 2 (class mixing proportion) \times 5 (class separation of location of knot) \times 3 (growth factor mean of slope of second phase) \times 2 (growth factor variance of slope of second phase) \times 2 (residual variance in observed variables)] was performed on only those replications that favored the correct model, the 2-class piecewise linear-linear LGMM, over the incorrect model, the 1-class piecewise linear-linear LGMM. The main effects of the manipulated factors and the interaction terms were reported and interpreted only when both the statistical significance ($p < .05$) and the practical significance ($\eta^2 \geq 0.06$) were achieved.

In all the ANOVA tables there were no main effects or interaction terms that satisfied both the statistical and the practical significance criteria. Thus, it can be concluded that the parameter bias and the variability index of parameter bias corresponding to only filtered replications are not systematically related to any of the

manipulated factors in the study. A summary of results obtained from the ANOVA tables are shown in Tables D3 and D4 in Appendix D.

4.4. Summary of Main Findings

The following are the main findings from the current study:

- For all the 360 cells, the parameter estimates for the intercept factor mean of the first phase, and the residual variance in observed variable were considered acceptable.
- For roughly 70% of the cells, the parameter estimates for slope factor mean of the second phase, and location of the knot were considered acceptable.
- For less than or equal to 39% of the cells, the parameter estimates for the variances of intercept and slope factors of the first phase, and slope factor of the second phase, were considered acceptable.
- The outcome measures, parameter bias and variability index of parameter bias, were not systematically related to any of the manipulated conditions in the design of the study.
- Among all the manipulated conditions, the condition of residual variance in observed variable had the strongest influence on both the model convergence rate and the model selection rate. Higher residual variance was associated with higher model convergence rate, and lower model selection rate. Lower residual variance was associated with lower model convergence rate, and higher model selection rate.

- Other manipulated conditions that influenced the model convergence rate and/or the model selection rate were growth factor mean of the slope of the second phase, growth factor variance of the slope of the second phase, and the class mixing proportion.
- The manipulated conditions that had no impact on either the model convergence rate or the model selection rate were sample size and class separation of location of the knot.

CHAPTER 5 – DISCUSSION

In the current study, a two-class piecewise linear-linear LGMM was developed, where the location of the knot is unknown. The model combines features of a piecewise linear-linear LGC model (where the location of the knot is unknown) with the ideas of latent class methods within the framework of SEM. The current research provided an in-depth analysis of accuracy of estimated model parameters when fitting a two-class piecewise linear-linear LGMM to data generated under different experimental conditions. Specifically, the performance of the two-class piecewise linear-linear LGMM was assessed under the manipulated conditions of sample size, class mixing proportion, class separation of location of knot, the mean of slope growth factor of the second phase, the variance of slope growth factor of the second phase, and residual variance of the observed variable. The outcome measures, parameter bias and variability index for parameter bias, were examined to see if they appeared to be a function of the manipulated conditions. Additionally, the effect of manipulated conditions on the percentage of properly converged replications for two-class piecewise linear-linear LGMMs across all the cells (i.e., 360) was analyzed.

Furthermore, the current study also addressed the issue of model misspecification. An analysis was conducted on model selection rate when fitting 1-, 2-, and 3-class piecewise linear-linear LGMMs to the data sets generated under different manipulated conditions using the 2-class piecewise linear-linear LGMM as the population model. The following sections include a summary of results, limitations of study, and methodological extensions.

5.1 Summary of Results

5.1.1 Model convergence rate

When fitting a two-class piecewise linear-linear LGMM under different manipulated conditions, the convergence rates to proper solutions seem to be most affected by the different levels of residual variance in observed variable condition. The large value of residual variance ($\sigma_{\epsilon}^2 = 5.0$) was associated with higher model convergence rate for the two-class piecewise linear-linear LGMM. This finding is not surprising because for convergence to be achieved, it is necessary to attain stationarity with respect to the estimation of parameters (i.e., the estimate for a parameter does not get any significantly better in subsequent iterations). The probability of finding a tenable value for a parameter perhaps increases with increase in observed variable variability, whether arising by residual or factor variance, because, spatially, the values that are sustainable take a larger range.

Other manipulated conditions that affected model convergence rate were the conditions of growth factor mean of slope of the second phase, growth factor variance of slope of the second phase, and class mixing proportion, when combined with different levels of residual variance condition. The conditions that had the highest model convergence rate had a combination of large value of residual variance ($\sigma_{\epsilon}^2 = 5.0$) and small value of slope2 variance ($\sigma_{\eta_{4k}}^2 = .20$). The conditions that had the worst model convergence rates had a combination of small value of residual variance ($\sigma_{\epsilon}^2 = 1.0$), either level 1 (i.e., low-low: 0.25, 0.25) or level 3 (i.e., high-high: 1.25, 1.25) of growth factor mean of slope of the second phase, and 50/50 class mixing proportion. The only

two manipulated conditions that did not have an effect on the model convergence rate were the conditions of sample size and class separation of location of the knot.

5.1.2 Parameter bias

The results related to the accuracy of parameter estimates suggest that the estimated variances of the growth factors, intercept and slope of the first phase, and slope of the second phase, were unsatisfactory. The percentage of cells with unacceptable values of the parameter estimates for the growth factor variances were more than 61%. The percentage of cells with unacceptable values of parameter estimates for the growth factor mean of the slope of the second phase 32.50% or less. The percentage of cells with unacceptable values for parameter estimates of the knot was nearly 30%. For the growth factor mean of the intercept of the first phase, and the residual variance in observed variable, none of the parameter estimates were considered unacceptable.

This finding is not that odd because the estimation of piecewise model, a type of partially nonlinear model, is known to be computationally intensive (piecewise model is a partially nonlinear model because the knot is a nonlinear parameter that does not have a random effect) (Cudeck & Klebe, 2002). Moreover, variances/covariances of the growth factors (random effects), especially the nonlinear, are notoriously problematic. The same explanation also applies to the estimation of location of the knot, where knot is a nonlinear parameter.

The results from 6-way ANOVA revealed that none of the manipulated conditions were systematically related to the outcome measures, parameter bias and variability index of parameter bias. Overall, the two-class piecewise linear-linear LGMM is a very

complex model and poor estimation of parameters reflect on the challenge with this model.

5.1.3 Model selection rate

The current research also investigated the possibility of extracting spurious latent classes. The effect of different manipulated conditions on the number of times the correct model (i.e., 2-class piecewise linear-linear LGMM) was preferred over incorrect models (i.e., 1- and 3- class piecewise linear-linear LGMMs) was analyzed. The manipulated condition of residual variance of observed variable had the strongest statistically significant effect on the model selection rate. The large value of residual variance ($\sigma_{\epsilon}^2 = 5.0$) was associated with lower model selection rate for the two-class piecewise linear-linear LGMM. This finding makes some sense because when fitting a model to data that has large value of residual variance, the sustainable parameter estimates could take on a wide range of values. This means the range could satisfy a wider variety of models, including both correctly and incorrectly specified models. Thus, a large value of residual variance may lead to lower selection rate.

The model selection rate was also affected by other manipulated conditions, such as growth factor mean of slope of the second phase, and class mixing proportion, when combined with different levels of residual variance condition. The conditions that had the best model convergence rates had a combination of small value of residual variance ($\sigma_{\epsilon}^2 = 1.0$), and either level 1 (i.e., low-low: 0.25, 0.25) or level 3 (i.e., high-high: 1.25, 1.25) of growth factor mean of slope of the second phase. The conditions that had the worst model convergence rate had a combination of large value of residual variance ($\sigma_{\epsilon}^2 = 5.0$) and 75/25 class mixing proportion. The only manipulated conditions that did

not have an effect on the model selection rate were the conditions of sample size, growth factor variance of slope of the second phase, and class separation of location of the knot.

5.2 Limitations of Study

The research design of the current study allowed for inferences about the performance of two-class piecewise linear-linear LGMM under a variety of manipulated conditions that were thought to be important to applied researchers. Nonetheless, as with any simulation study, there are an infinite number of combinations of manipulated conditions that could have been analyzed. For example, this research did not incorporate observed variables with non-normal distributions, or different types of residual covariance structures, such as first-order auto-regressive covariance structure or Toeplitz covariance structure.

Another limitation of the study is that it did not incorporate any observed or latent covariates. Inclusion of covariates in a growth mixture model may help in improving model convergence and latent class membership (Lubke & Muthén, 2007).

Furthermore, the results were based on only those replications that converged to a proper solution, which means that the results should be considered the upper bounds. The replications that did not converge to a proper solution were discarded. Hence, the generalizability of the findings for the cells with relatively low convergence rate are necessarily impacted (Hamilton, 2009).

Lastly, the current study was limited to one type of model selection index. While BIC is a popular index of model selection in applied research, there are conflicting results about its effectiveness when sample size is large. That is, Tolvanen (2008) recommended

using BIC with small sample sizes and aBIC (adjusted BIC) with large sample sizes. The borderline between small and large sample size was suggested to be about 500.

5.3 Recommendations for Applied Researchers

In applied settings it may not be obvious for researchers to know which model to pick that not only fits the data best, but also makes sense substantively. A couple of informed steps can help applied researchers to select and fit the right model to data. Below is the general guideline that a researcher may follow:

First, it is important for a researcher to visually see the data via graphs, such as spaghetti plots, to get a good sense of the functional form of individual's growth over time, unless he or she has a theory that dictates the functional form of individual's growth. Furthermore, the visual inspection of the graphs will also give an idea about whether or not the overall functional form appears to be made up of different segments. That is, whether or not there is a piecewise change in the individual's trajectory over time. Additionally, there are tests for determining empirically whether the within-subject functional form truly is piecewise rather than some other smooth function. The parameterization allows testing the homogeneity of slopes of segment 1 and 2. If the test is found to be statistically significant, it would imply that the two slopes are not equal. This result would indicate the existence of within-subject piecewise function.

Second, a piecewise LGMM can be considered to fit data if a researcher believes the observed sample is made up of multiple population distributions with distinct characteristics and that the functional form in each latent class has a disjuncture. It may not, however, be advisable to straight away fit a piecewise LGMM without thinking

about the issues related with LGMMs. That is, LGMMs, in general, have a hard time to converge to a stable solution. Moreover, LGMMs also have the problem of converging to a local maximum of the likelihood. Ideally, it is recommended that population values should be provided as starting values of the parameters to be estimated to minimize convergence problems.

In applied settings the population values are not known, however. A useful way for applied researchers to minimize convergence problems is to use previous research to estimate appropriate starting values. Another way to minimize convergence related issues is to estimate the model parts separately to obtain appropriate starting values for the full model. For example, a 1-class piecewise LGC model can be fit to data and then later, the estimated values of model parameters can be used as starting values of the LGMM's parameters to be estimated. Another alternative step can be to fit a latent class growth model in which individuals within a class are treated as homogeneous with respect to their development, that is, the within-class between-subject variability is suppressed. The estimated parameter values corresponding to the mean structure and residual variance will give some indication for the starting values of the mean structure and residual variance for the LGMM.

Furthermore, if posterior class membership probability is of interest to applied researchers, a potential way of improving it is to add class-predicting covariates to the model (Lubke & Muthén, 2007). There is a caveat here, however. The class-predicting covariates would make the model more complex, and so convergence problems may result.

5.4 Methodological Extensions

The current study provided a preliminary evaluation of the viability of two-class piecewise linear-linear LGMM, where the location of knot is unknown and fixed for all individual cases. The issues related to the estimation of model parameters needs to be investigated more. It may be worthwhile exploring in future studies how the model fits when using MCMC estimation method and compare the results obtained from ML estimation via EM algorithm. Once the issues related with the estimation of two-class piecewise linear-linear LGMM are resolved some interesting extensions can be considered, such as the following:

It will be very interesting to extend this model to include a random effect for knot, i.e., each individual case is allowed to have its own location of knot, and analyze problems with respect to the estimation procedure. Another interesting extension can be to incorporate observed and latent covariates to predict the location of knot.

In this study it was assumed that both the classes have the same functional form of piecewise change process, i.e., linear-linear. It will be interesting to incorporate different functional form of piecewise change process for each latent class, for example one class has a linear-linear piecewise change process and the other class has linear-exponential piecewise change process.

In sum, the research reported here seem to indicate that a two-class piecewise linear-linear LGMM is a very complex model and that there are issues related to the estimation of model parameters. At the same time, it also seems to be a useful and flexible model in the area of educational research where most often the interest of researchers is centered on student academic progress or changes in attitude and affect, for

example researchers may be interested in studying the effectiveness of treatment/intervention on students' academic progress where the population of students is composed of two or more latent groups. The utility of two-class piecewise linear-linear LGMM is that it allows researchers to specify each developmental phase to conform to a particular form of the overall change process within each latent class.

APPENDIX – A

The Procedure of Reparameterization:

Based on Harring et al. (2006) paper, the three combinations of the slopes and intercepts is given as:

$$\omega_1 = \frac{1}{2}(\eta_{1i} + \eta_{3i}) \quad \omega_2 = \frac{1}{2}(\eta_{2i} + \eta_{4i}) \quad \omega_3 = \frac{1}{2}(\eta_{4i} - \eta_{2i}) \quad (1)$$

When the mean of slope of second phase, α_4 , is greater than the mean of slope of first phase, α_2 , (i.e., $\alpha_4 > \alpha_2$), the reparameterized model can be written as the maximum of the two segments (Harring et al., 2006). That is,

$$f = \max(l_1(x), l_2(x))$$

A convenient form of the max function can be written as:

$$\max(u, v) = \frac{1}{2}(u + v + \sqrt{(u - v)^2}) \quad (2)$$

Substituting the segments $l_1(x)$ and $l_2(x)$ into the above equation gives:

$$\begin{aligned} f &= \max(l_1, l_2) \\ &= \frac{1}{2}(l_1 + l_2 + \sqrt{(l_1 - l_2)^2}) \\ &= \frac{1}{2}\{\eta_{1i} + \eta_{2i}t_{ji} + \eta_{3i} + \eta_{4i}t_{ji} + \sqrt{[(\eta_{1i} + \eta_{2i}t_{ji}) - (\eta_{3i} + \eta_{4i}t_{ji})]^2}\} \\ &= \frac{1}{2}\{(\eta_{1i} + \eta_{3i}) + (\eta_{2i} + \eta_{4i})t_{ji} + \sqrt{[(\eta_{2i} - \eta_{4i})t_{ji} + (\eta_{1i} - \eta_{3i})]^2}\} \end{aligned}$$

Because the segments join at the knot (i.e., when $t_{ji} = \gamma$, $\eta_{1i} + \eta_{2i}\gamma = \eta_{3i} + \eta_{4i}\gamma$),

therefore

$\eta_{1i} - \eta_{3i} = (\eta_{4i} - \eta_{2i})\gamma$. Consequently,

$$\begin{aligned} f &= \frac{1}{2}\{(\eta_{1i} + \eta_{3i}) + (\eta_{2i} + \eta_{4i})t_{ji} + \sqrt{[(\eta_{2i} - \eta_{4i})t_{ji} + (\eta_{4i} - \eta_{2i})\gamma]^2}\} \\ &= \frac{1}{2}\{(\eta_{1i} + \eta_{3i}) + (\eta_{2i} + \eta_{4i})t_{ji} + \sqrt{[(\eta_{2i} - \eta_{4i})(t_{ji} - \gamma)]^2}\} \end{aligned}$$

$$= \frac{1}{2} \{ (\eta_{1i} + \eta_{3i}) + (\eta_{2i} + \eta_{4i})t_{ji} - ((\eta_{4i} - \eta_{2i})\sqrt{(t_{ji} - \gamma)^2}) \}$$

Using the terms in Equation (1) gives the reparameterized piecewise linear-linear LGC model with unknown knot location as:

$$f = \omega_1 + \omega_2 t_{ji} - \omega_3 \sqrt{(t_{ji} - \gamma)^2} \quad (3)$$

The model in Equation (3) is the one that will be fit using standard SEM software. The coefficients of the original model in Equation (2.24) can be reconstructed from $\hat{\omega}_1$, $\hat{\omega}_2$, $\hat{\omega}_3$ and $\hat{\gamma}$ as:

$$\hat{\eta}_{1i} = \hat{\omega}_1 - \hat{\omega}_3 \hat{\gamma} \quad \hat{\eta}_{2i} = \hat{\omega}_2 + \hat{\omega}_3 \quad \hat{\eta}_{4i} = \hat{\omega}_2 - \hat{\omega}_3$$

Note that when the mean of slope of second phase, α_4 , is less than the mean of slope of first phase, α_4 , (i.e., $\alpha_4 < \alpha_2$), the model in Equation (2.24) can be rewritten as the minimum of the two segments (Harring et al., 2006). That is,

$$f = \min(l_1(x), l_2(x))$$

The only change in procedure is that (3) is replaced by

$$f = \omega_1 + \omega_2 t_{ji} + \omega_3 \sqrt{(t_{ji} - \gamma)^2}$$

APPENDIX – B

The Multivariate Delta Method of Transformation:

The delta method transforms the estimated variances of ω_1 , ω_2 , and ω_3 (i.e., $\hat{\sigma}_{\omega_1}^2$, $\hat{\sigma}_{\omega_2}^2$ and $\hat{\sigma}_{\omega_3}^2$) back to the respective variances of η_{1i} , η_{2i} , and η_{4i} (i.e., $\hat{\sigma}_{\eta_{1i}}^2$, $\hat{\sigma}_{\eta_{2i}}^2$ and $\hat{\sigma}_{\eta_{4i}}^2$) in the following way:

$$\hat{\eta}_{1i} = \hat{\omega}_1 - \hat{\omega}_3 \hat{\gamma} = g_1(\omega)$$

$$\hat{\eta}_{2i} = \hat{\omega}_2 + \hat{\omega}_3 = g_2(\omega)$$

$$\hat{\eta}_{4i} = \hat{\omega}_2 - \hat{\omega}_3 = g_3(\omega)$$

- 1) $Var(\hat{\sigma}_{\eta_{1i}}^2) = \mathbf{D}^T \mathbf{\Sigma} \mathbf{D}$, where \mathbf{D} is the matrix of partial derivatives of g_1 with respect to ω_1 , ω_2 , and ω_3 .

$$\begin{aligned} &= \begin{bmatrix} \frac{\partial(\eta_{1i})}{\partial(\omega_1)} & \frac{\partial(\eta_{1i})}{\partial(\omega_2)} & \frac{\partial(\eta_{1i})}{\partial(\omega_3)} \end{bmatrix} \cdot \begin{bmatrix} \hat{\sigma}_{\omega_1}^2 & & \\ \hat{\sigma}_{\omega_2, \omega_1} & \hat{\sigma}_{\omega_2}^2 & \\ \hat{\sigma}_{\omega_3, \omega_1} & \hat{\sigma}_{\omega_3, \omega_1} & \hat{\sigma}_{\omega_3}^2 \end{bmatrix} \cdot \begin{bmatrix} \frac{\partial(\eta_{1i})}{\partial(\omega_1)} \\ \frac{\partial(\eta_{1i})}{\partial(\omega_2)} \\ \frac{\partial(\eta_{1i})}{\partial(\omega_3)} \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & -\hat{\gamma} \end{bmatrix} \cdot \begin{bmatrix} \hat{\sigma}_{\omega_1}^2 & & \\ \hat{\sigma}_{\omega_2, \omega_1} & \hat{\sigma}_{\omega_2}^2 & \\ \hat{\sigma}_{\omega_3, \omega_1} & \hat{\sigma}_{\omega_3, \omega_1} & \hat{\sigma}_{\omega_3}^2 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ -\hat{\gamma} \end{bmatrix} \end{aligned}$$

- 2) $Var(\hat{\sigma}_{\eta_{2i}}^2) = \mathbf{D}^T \mathbf{\Sigma} \mathbf{D}$, where \mathbf{D} is the matrix of partial derivatives of g_2 with respect to ω_1 , ω_2 , and ω_3 .

$$= \begin{bmatrix} \frac{\partial(\eta_{2i})}{\partial(\omega_1)} & \frac{\partial(\eta_{2i})}{\partial(\omega_2)} & \frac{\partial(\eta_{2i})}{\partial(\omega_3)} \end{bmatrix} \cdot \begin{bmatrix} \hat{\sigma}_{\omega_1}^2 & & \\ \hat{\sigma}_{\omega_2, \omega_1} & \hat{\sigma}_{\omega_2}^2 & \\ \hat{\sigma}_{\omega_3, \omega_1} & \hat{\sigma}_{\omega_3, \omega_1} & \hat{\sigma}_{\omega_3}^2 \end{bmatrix} \cdot \begin{bmatrix} \frac{\partial(\eta_{2i})}{\partial(\omega_1)} \\ \frac{\partial(\eta_{2i})}{\partial(\omega_2)} \\ \frac{\partial(\eta_{2i})}{\partial(\omega_3)} \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} \hat{\sigma}_{\omega_1}^2 & & \\ \hat{\sigma}_{\omega_2, \omega_1} & \hat{\sigma}_{\omega_2}^2 & \\ \hat{\sigma}_{\omega_3, \omega_1} & \hat{\sigma}_{\omega_3, \omega_1} & \hat{\sigma}_{\omega_3}^2 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

3) $Var(\hat{\sigma}_{\eta_{4i}}^2) = \mathbf{D}^T \mathbf{\Sigma} \mathbf{D}$, where \mathbf{D} is the matrix of partial derivatives of g_3 with respect to ω_1 , ω_2 , and ω_3 .

$$= \begin{bmatrix} \frac{\partial(\eta_{4i})}{\partial(\omega_1)} & \frac{\partial(\eta_{4i})}{\partial(\omega_2)} & \frac{\partial(\eta_{4i})}{\partial(\omega_3)} \end{bmatrix} \cdot \begin{bmatrix} \hat{\sigma}_{\omega_1}^2 & & \\ \hat{\sigma}_{\omega_2, \omega_1} & \hat{\sigma}_{\omega_2}^2 & \\ \hat{\sigma}_{\omega_3, \omega_1} & \hat{\sigma}_{\omega_3, \omega_1} & \hat{\sigma}_{\omega_3}^2 \end{bmatrix} \cdot \begin{bmatrix} \frac{\partial(\eta_{4i})}{\partial(\omega_1)} \\ \frac{\partial(\eta_{4i})}{\partial(\omega_2)} \\ \frac{\partial(\eta_{4i})}{\partial(\omega_3)} \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 1 & -1 \end{bmatrix} \cdot \begin{bmatrix} \hat{\sigma}_{\omega_1}^2 & & \\ \hat{\sigma}_{\omega_2, \omega_1} & \hat{\sigma}_{\omega_2}^2 & \\ \hat{\sigma}_{\omega_3, \omega_1} & \hat{\sigma}_{\omega_3, \omega_1} & \hat{\sigma}_{\omega_3}^2 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$$

APPENDIX – C

Results from the Pilot Study for Sample Size Selection:

Table C1.

Proportion of successful replications.

	$N = 200$	$N = 400$	$N = 700$	$N = 1000$	$N = 2000$
Cell 1	0.85	0.91	0.95	0.98	1.00
Cell 2	0.83	0.89	0.91	0.91	0.95
Cell 3	0.88	0.89	0.91	0.95	0.99
Cell 4	0.93	0.97	0.97	0.99	1.00
Cell 5	0.89	0.91	0.87	0.92	0.90
Cell 6	0.88	0.94	0.96	0.95	0.97

Table C2.

Average parameter bias and variance around the average parameter bias for $n = 200$.

		Cell 1	Cell 2	Cell 3	Cell 4	Cell 5	Cell 6
Class 1	Intercept 1	-0.04 (0.322)	-0.072 (0.131)	-0.007 (0.258)	-0.067 (0.18)	0.002 (0.099)	-0.046 (0.172)
	Slope 1	-0.067 (0.248)	-0.04 (0.076)	-0.104 (0.166)	0.054 (0.14)	0.059 (0.063)	0.025 (0.21)
	Slope 2	0.126 (0.557)	11.559 (10719.29)	-10.162 (9184.315)	0.246 (0.646)	0.106 (0.818)	0.241 (0.735)
	Var(Intercept 1)	0.242 (0.327)	0.185 (0.276)	0.030 (0.238)	0.389 (0.460)	0.112 (0.312)	0.218 (2.578)
	Var(Slope 1)	0.248 (0.008)	0.082 (0.005)	0.227 (0.007)	0.266 (0.009)	0.063 (0.004)	0.191 (0.008)
	Var(Slope 2)	-0.552 (0.008)	-0.718 (0.005)	-0.573 (0.007)	-0.534 (0.009)	-0.737 (0.004)	-0.609 (0.008)
	Knot	-0.093 (0.642)	0.116 (1.066)	0.048 (1.257)	-0.099 (0.422)	0.268 (0.980)	0.621 (2.001)
Class 2	Intercept 1	0.048 (0.259)	0.032 (0.161)	-0.047 (0.434)	0.002 (0.269)	-0.053 (0.116)	0.015 (0.282)
	Slope 1	0.029 (0.124)	0.034 (0.117)	0.528 (25.566)	-0.078 (0.149)	0.107 (0.113)	0.093 (0.106)
	Slope 2	-0.068 (0.571)	-0.129 (0.755)	-15.356 (20842.76)	-0.591 (0.765)	10.693 (6171.148)	-12.828 (12342.42)
	Var(Intercept 1)	0.260 (0.317)	0.137 (0.263)	0.121 (0.286)	0.497 (0.530)	0.103 (0.306)	0.280 (3.307)
	Var(Slope 1)	0.248 (0.008)	0.082 (0.005)	0.227 (0.007)	0.266 (0.009)	0.063 (0.004)	0.191 (0.008)
	Var(Slope 2)	-0.552 (0.008)	-0.718 (0.005)	-0.573 (0.007)	-0.534 (0.009)	-0.737 (0.004)	-0.609 (0.008)
	Knot	-1.110 (0.227)	-2.029 (0.538)	-2.899 (1.378)	-0.954 (0.426)	-1.656 (1.335)	-2.347 (1.862)

Table C3.

Average parameter bias and variance around the average parameter bias for $n=400$.

		Cell 1	Cell 2	Cell 3	Cell 4	Cell 5	Cell 6
Class 1	Intercept 1	-0.029 (0.217)	0.02 (0.092)	0.000 (0.293)	-0.008 (0.232)	-0.036 (0.082)	0.034 (0.195)
	Slope 1	-0.072 (0.094)	-0.006 (0.058)	-0.034 (0.082)	-0.023 (0.047)	0.017 (0.052)	0.000 (0.064)
	Slope 2	0.091 (0.45)	-0.023 (0.469)	0.006 (0.365)	0.272 (0.627)	0.251 (0.68)	0.081 (0.512)
	Var(Intercept 1)	0.237 (0.089)	0.177 (0.123)	0.091 (0.086)	0.483 (0.281)	0.192 (0.221)	0.141 (1.036)
	Var(Slope 1)	0.262 (0.004)	0.089 (0.003)	0.254 (0.004)	0.279 (0.004)	0.067 (0.003)	0.204 (0.006)
	Var(Slope 2)	-0.538 (0.004)	-0.711 (0.003)	-0.546 (0.004)	-0.521 (0.004)	-0.733 (0.003)	-0.596 (0.006)
	Knot	-0.174 (0.116)	-0.064 (0.445)	-0.141 (0.208)	-0.096 (0.231)	0.208 (0.523)	0.396 (0.895)
	Intercept 1	0.016 (0.289)	-0.049 (0.147)	-0.014 (0.314)	-0.051 (0.204)	0.025 (0.071)	-0.045 (0.255)
	Slope 1	0.049 (0.06)	-0.016 (0.05)	-0.018 (0.131)	-0.003 (0.043)	0.059 (0.046)	0.094 (0.076)
	Slope 2	-0.03 (0.527)	0.095 (0.629)	0.014 (0.465)	-0.587 (0.689)	-0.976 (0.74)	-6.708 (3291.963)
Class 2	Var(Intercept 1)	0.296 (0.113)	0.242 (0.147)	0.073 (0.104)	0.616 (0.311)	0.157 (0.156)	0.344 (1.892)
	Var(Slope 1)	0.262 (0.004)	0.089 (0.003)	0.254 (0.004)	0.279 (0.004)	0.067 (0.003)	0.204 (0.006)
	Var(Slope 2)	-0.538 (0.004)	-0.711 (0.003)	-0.546 (0.004)	-0.521 (0.004)	-0.733 (0.003)	-0.596 (0.006)
	Knot	-1.101 (0.121)	-2.008 (0.226)	-3.178 (0.179)	-0.978 (0.222)	-1.817 (0.521)	-2.442 (1.568)

Table C4.

Average parameter bias and variance around the average parameter bias for $n=700$.

		Cell 1	Cell 2	Cell 3	Cell 4	Cell 5	Cell 6
Class 1	Intercept 1	-0.013 (0.196)	0.014 (0.06)	-0.003 (0.244)	-0.022 (0.174)	-0.034 (0.055)	0.037 (0.171)
	Slope 1	-0.026 (0.075)	0.029 (0.049)	-0.037 (0.075)	0.007 (0.037)	0.024 (0.04)	0.023 (0.056)
	Slope 2	0.044 (0.447)	-0.1 (0.491)	-0.022 (0.295)	0.299 (0.63)	0.23 (0.622)	0.039 (0.466)
	Var(Intercept 1)	0.286 (0.045))	0.218 (0.076)	0.087 (0.051)	0.537 (0.243)	0.169 (0.105)	0.111 (0.146)
	Var(Slope 1)	0.259 (0.002)	0.085 (0.002)	0.259 (0.002)	0.284 (0.002)	0.067 (0.002)	0.205 (0.003)
	Var(Slope 2)	-0.541 (0.002)	-0.715 (0.002)	-0.541 (0.002)	-0.516 (0.002)	-0.733 (0.002)	-0.595 (0.003)
	Knot	-0.176 (0.070)	-0.1 (0.177)	-0.252 (0.068)	-0.091 (0.237)	0.187 (0.460)	0.224 (0.542)
Class 2	Intercept 1	-0.011 (0.212)	-0.057 (0.088)	-0.03 (0.276)	-0.082 (0.182)	0.02 (0.61)	-0.024 (0.191)
	Slope 1	-0.021 (0.079)	-0.051 (0.057)	-0.059 (0.094)	-0.034 (0.041)	0.047 (0.039)	0.061 (0.069)
	Slope 2	0.019 (0.49)	0.195 (0.515)	0.06 (0.405)	-0.567 (0.63)	-0.942 (0.708)	-0.847 (0.554)
	Var(Intercept 1)	0.310 (0.092)	0.229 (0.083)	0.141 (0.073)	0.620 (0.299)	0.178 (0.114)	0.098 (0.129)
	Var(Slope 1)	0.259 (0.002)	0.085 (0.002)	0.259 (0.002)	0.284 (0.002)	0.067 (0.002)	0.205 (0.003)
	Var(Slope 2)	-0.541 (0.002)	-0.715 (0.002)	-0.541 (0.002)	-0.516 (0.002)	-0.733 (0.002)	-0.595 (0.003)
	Knot	-1.163 (0.072)	-2.064 (0.207)	-3.185 (0.072)	-1.030 (0.201)	-1.801 (0.359)	-2.658 (0.856)

Table C5.

Average parameter bias and variance around the average parameter bias for $n=1000$.

		Cell 1	Cell 2	Cell 3	Cell 4	Cell 5	Cell 6
Class 1	Intercept 1	-0.002 (0.192)	-0.018 (0.071)	-0.046 (0.23)	0.029 (0.145)	-0.034 (0.044)	0.051 (0.121)
	Slope 1	-0.024 (0.07)	0.009 (0.051)	-0.065 (0.081)	0.023 (0.033)	0.005 (0.039)	0.014 (0.064)
	Slope 2	0.025 (0.453)	-0.04 (0.482)	0.053 (0.329)	0.193 (0.548)	0.248 (0.628)	0.037 (0.403)
	Var(Intercept 1)	0.299 (0.045)	0.261 (0.069)	0.130 (0.053)	0.462 (0.214)	0.241 (0.091)	0.108 (0.101)
	Var(Slope 1)	0.262 (0.002)	0.092 (0.001)	0.262 (0.002)	0.279 (0.001)	0.068 (0.001)	0.211 (0.002)
	Var(Slope 2)	-0.538 (0.002)	-0.708 (0.001)	-0.538 (0.002)	-0.521 (0.001)	-0.732 (0.001)	-0.589 (0.002)
	Knot	-0.184 (0.051)	-0.134 (0.158)	-0.257 (0.045)	-0.161 (0.187)	0.165 (0.331)	0.035 (0.365)
Class 2	Intercept 1	-0.005 (0.2)	-1.314 (9.012)	-0.027 (0.07)	-0.121 (0.174)	0.019 (0.049)	0.001 (0.186)
	Slope 1	-0.025 (0.065)	0.369 (0.522)	-0.025 (0.048)	-0.039 (0.038)	0.069 (0.036)	0.038 (0.055)
	Slope 2	0.01 (0.471)	-0.275 (0.048)	0.119 (0.522)	-0.499 (0.648)	-0.984 (0.667)	-0.88 (0.563)
	Var(Intercept 1)	0.340 (0.065)	0.290 (0.065)	0.152 (0.042)	0.684 (0.204)	0.204 (0.076)	0.234 (0.127)
	Var(Slope 1)	0.262 (0.002)	0.092 (0.001)	0.262 (0.002)	0.279 (0.001)	0.068 (0.001)	0.211 (0.002)
	Var(Slope 2)	-0.538 (0.002)	-0.708 (0.001)	-0.538 (0.002)	-0.521 (0.001)	-0.732 (0.001)	-0.589 (0.002)
	Knot	-1.143 (0.049)	-2.066 (0.088)	-3.224 (0.040)	-0.944 (0.174)	-1.876 (0.301)	-2.721 (0.554)

Table C6.

Average parameter bias and variance around the average parameter bias for $n=2000$.

		Cell 1	Cell 2	Cell 3	Cell 4	Cell 5	Cell 6
Class 1	Intercept 1	0.009 (0.179)	-0.019 (0.059)	0.109 (0.241)	0.012 (0.148)	-0.023 (0.035)	0.128 (0.107)
	Slope 1	-0.014 (0.066)	0.007 (0.048)	0.017 (0.058)	0.033 (0.032)	0.005 (0.038)	0.064 (0.047)
	Slope 2	-0.017 (0.445)	0.03 (0.51)	-0.142 (0.296)	0.211 (0.596)	0.255 (0.621)	-0.065 (0.419)
	Var(Intercept 1)	0.305 (0.024)	0.303 (0.028)	0.130 (0.018)	0.507 (0.208)	0.283 (0.054)	0.112 (0.065)
	Var(Slope 1)	0.262 (0.001)	0.092 (0.001)	0.262 (0.001)	0.282 (0.001)	0.072 (0.000)	0.209 (0.001)
	Var(Slope 2)	-0.538 (0.001)	-0.708 (0.001)	-0.538 (0.001)	-0.518 (0.001)	-0.728 (0.000)	-0.591 (0.001)
	Knot	-0.176 (0.031)	-0.066 (0.070)	-0.239 (0.025)	-0.106 (0.188)	0.177 (0.168)	0.088 (0.276)
	Intercept 1	-0.012 (0.183)	-0.02 (0.056)	-0.128 (0.215)	-0.104 (0.155)	0.025 (0.033)	-0.085 (0.141)
	Slope 1	-0.026 (0.063)	-0.016 (0.048)	-0.118 (0.071)	-0.038 (0.031)	0.063 (0.038)	0.013 (0.048)
	Slope 2	0.043 (0.453)	0.036 (0.486)	0.145 (0.322)	-0.517 (0.602)	-0.999 (0.614)	-0.782 (0.475)
Class 2	Var(Intercept 1)	0.333 (0.029)	0.288 (0.025)	0.156 (0.020)	0.660 (0.182)	0.230 (0.047)	0.250 (0.090)
	Var(Slope 1)	0.262 (0.001)	0.092 (0.001)	0.262 (0.001)	0.282 (0.001)	0.072 (0.000)	0.209 (0.001)
	Var(Slope 2)	-0.538 (0.001)	-0.708 (0.001)	-0.538 (0.001)	-0.518 (0.001)	-0.728 (0.000)	-0.591 (0.001)
	Knot	-1.142 (0.042)	-2.113 (0.064)	-3.200 (0.032)	-0.956 (0.169)	-1.968 (0.216)	-2.715 (0.349)

Results from the Pilot Study for Asymptotic Behavior of Model Parameters for a 2-Class Piecewise Linear-Linear LGMM:

Table C7.

Population values for the data generation of the 2-class piecewise linear-linear LGMM.

	Class 1	Class 2
Growth Factor Means of Intercept and Slope of First Phase		
• Intercept 1 (α_{1k})	2.0	2.0
• Slope 1 (α_{2k})	0.0	0.0
Growth Factor Mean of Slope of Second Phase (α_{4k})	0.25	1.25
Growth Factor Variances of Intercept and Slope of First Phase		
• Intercept 1 ($\sigma_{\eta_{1k}}^2$)	1.0	1.0
• Slope 1 ($\sigma_{\eta_{2k}}^2$)	0.2	0.2
Growth Factor Variance of Slope of Second Phase ($\sigma_{\eta_{4k}}^2$)	0.2	0.2
Growth Factor Covariances		
• Intercepts and slopes	0.0	0.0
Class Separation of Location of Knot (γ_k)	2.0	4.0
Residual Variance of Observed Variables (σ_{ϵ}^2)	1.0	1.0

Table C8.

Estimated original model parameters of the 2-class piecewise linear-linear LGMM.

	Class 1	Class 2
Growth Factor Means of Intercept and Slope of First Phase		
• Intercept 1 (α_{1k})	1.998	2.004
• Slope 1 (α_{2k})	0.004	-0.01
Growth Factor Mean of Slope of Second Phase (α_{4k})	0.25	1.256
Growth Factor Variances of Intercept and Slope of First Phase		
• Intercept 1 ($\sigma_{\eta_{1k}}^2$)	1.026	1.003
• Slope 1 ($\sigma_{\eta_{2k}}^2$)	0.205	0.2
Growth Factor Variance of Slope of Second Phase ($\sigma_{\eta_{4k}}^2$)	0.201	0.2
Class Separation of Location of Knot (γ_k)	2.008	4.001
Residual Variance of Observed Variables (σ_{ϵ}^2)	0.998	0.994

APPENDIX – D

Table D1.

Tests of between-subject effects on parameter bias.

		Statistical and Practical Significance	Manipulated Factors					
			Sample Size, n	Class Mixing Proportion, φ_k	Class Separation Location of Knot, γ_k	Mean of Slope of Second Phase, α_{4k}	Variance of Slope of Second Phase, $\sigma_{\eta_{4k}}^2$	Residual Variance σ_{ϵ}^2
Parameter Bias (Class 1)	Bias_ α_{11}	p-value =	0.439	0.329	0.326	0.448	0.491	0.150
		partial η^2 =	0.000	0.000	0.000	0.000	0.000	0.000
	Bias_ α_{21}	p-value =	0.957	0.183	0.577	0.274	0.961	0.354
		partial η^2 =	0.000	0.000	0.000	0.000	0.000	0.000
	Bias_ α_{41}	p-value =	0.519	0.299	0.427	0.179	0.796	0.363
		partial η^2 =	0.000	0.000	0.000	0.000	0.000	0.000
	Bias_ γ_1	p-value =	0.027	0.000	0.000	0.000	0.000	0.000
		partial η^2 =	0.000	0.001	0.019	0.001	0.002	0.004
	Bias_ $\sigma_{\eta_{11}}^2$	p-value =	0.020	0.019	0.037	0.000	0.000	0.000
		partial η^2 =	0.000	0.000	0.000	0.001	0.003	0.002
	Bias_ $\sigma_{\eta_{21}}^2$	p-value =	0.003	0.415	0.467	0.000	0.000	0.000
		partial η^2 =	0.000	0.000	0.000	0.002	0.003	0.001
Bias_ $\sigma_{\eta_{41}}^2$	p-value =	0.000	0.011	0.007	0.000	0.000	0.000	
	partial η^2 =	0.001	0.000	0.000	0.003	0.02	0.003	
Bias_ $\sigma_{\epsilon_1}^2$	p-value =	0.220	0.423	0.002	0.084	0.000	0.000	
	partial η^2 =	0.000	0.000	0.001	0.000	0.001	0.002	
Parameter Bias (Class 2)	Bias_ α_{12}	p-value =	0.281	0.056	0.006	0.450	0.375	0.000
		partial η^2 =	0.000	0.000	0.000	0.000	0.000	0.001
	Bias_ α_{22}	p-value =	0.262	0.711	0.255	0.251	0.832	0.409
		partial η^2 =	0.000	0.000	0.000	0.000	0.000	0.000
	Bias_ α_{42}	p-value =	0.360	0.241	0.925	0.031	0.469	0.737
		partial η^2 =	0.000	0.000	0.000	0.000	0.000	0.000
	Bias_ γ_2	p-value =	0.036	0.000	0.000	0.000	0.021	0.000
		partial η^2 =	0.000	0.002	0.03	0.004	0.000	0.004
	Bias_ $\sigma_{\eta_{12}}^2$	p-value =	0.532	0.918	0.001	0.000	0.000	0.000
		partial η^2 =	0.000	0.000	0.001	0.002	0.004	0.001
	Bias_ $\sigma_{\eta_{22}}^2$	p-value =	0.113	0.131	0.845	0.045	0.000	0.038
		partial η^2 =	0.000	0.000	0.000	0.000	0.000	0.000
Bias_ $\sigma_{\eta_{42}}^2$	p-value =	0.158	0.265	0.654	0.000	0.000	0.000	
	partial η^2 =	0.000	0.000	0.000	0.001	0.007	0.001	
Bias_ $\sigma_{\epsilon_2}^2$	p-value =	0.085	0.377	0.000	0.169	0.051	0.000	
	partial η^2 =	0.000	0.000	0.001	0.000	0.000	0.001	

Table D2.

Tests of between-subject effects on variability index for parameter bias.

		Statistical and Practical Significance	Manipulated Factors					
			Sample Size, n	Class Mixing Proportion, φ_k	Class Separation Location of Knot, γ_k	Mean of Slope of Second Phase, α_{4k}	Variance of Slope of Second Phase, $\sigma_{\eta_{4k}}^2$	Residual Variance σ_{ϵ}^2
Variability Index (Class 1)	Bias_ α_{11}	p-value =	0.469	0.329	0.383	0.523	0.477	0.384
		partial η^2 =	0.000	0.000	0.000	0.000	0.000	0.000
	Bias_ α_{21}	p-value =	0.556	0.374	0.752	0.598	0.884	0.131
		partial η^2 =	0.000	0.000	0.000	0.000	0.000	0.000
	Bias_ α_{41}	p-value =	0.468	0.364	0.551	0.518	0.353	0.352
		partial η^2 =	0.000	0.000	0.000	0.000	0.000	0.000
	Bias_ γ_1	p-value =	0.001	0.002	0.000	0.000	0.000	0.000
		partial η^2 =	0.000	0.000	0.003	0.002	0.003	0.005
	Bias_ $\sigma_{\eta_{11}}^2$	p-value =	0.668	0.042	0.382	0.206	0.551	0.000
		partial η^2 =	0.000	0.000	0.000	0.000	0.000	0.000
	Bias_ $\sigma_{\eta_{21}}^2$	p-value =	0.414	0.783	0.829	0.088	0.752	0.063
		partial η^2 =	0.000	0.000	0.000	0.000	0.000	0.000
Bias_ $\sigma_{\eta_{41}}^2$	p-value =	0.618	0.097	0.387	0.050	0.736	0.000	
	partial η^2 =	0.000	0.000	0.000	0.000	0.000	0.001	
Bias_ $\sigma_{\epsilon_1}^2$	p-value =	0.000	0.000	0.000	0.001	0.000	0.000	
	partial η^2 =	0.002	0.002	0.004	0.000	0.000	0.039	
Variability Index (Class 2)	Bias_ α_{12}	p-value =	0.491	0.318	0.497	0.534	0.467	0.240
		partial η^2 =	0.000	0.000	0.000	0.000	0.000	0.000
	Bias_ α_{22}	p-value =	0.378	0.488	0.388	0.546	0.486	0.321
		partial η^2 =	0.000	0.000	0.000	0.000	0.000	0.000
	Bias_ α_{42}	p-value =	0.570	0.480	0.428	0.523	0.262	0.541
		partial η^2 =	0.000	0.000	0.000	0.000	0.000	0.000
	Bias_ γ_2	p-value =	0.001	0.005	0.000	0.000	0.000	0.000
		partial η^2 =	0.000	0.002	0.002	0.002	0.001	0.004
	Bias_ $\sigma_{\eta_{12}}^2$	p-value =	0.583	0.669	0.867	0.160	0.215	0.018
		partial η^2 =	0.000	0.000	0.000	0.000	0.000	0.000
	Bias_ $\sigma_{\eta_{22}}^2$	p-value =	0.150	0.168	0.670	0.183	0.219	0.164
		partial η^2 =	0.000	0.000	0.000	0.000	0.000	0.000
Bias_ $\sigma_{\eta_{42}}^2$	p-value =	0.177	0.176	0.688	0.102	0.327	0.085	
	partial η^2 =	0.000	0.000	0.000	0.000	0.000	0.000	
Bias_ $\sigma_{\epsilon_2}^2$	p-value =	0.000	0.000	0.000	0.000	0.000	0.000	
	partial η^2 =	0.001	0.002	0.005	0.004	0.001	0.037	

NOTE that in the tests of between-subject effects on parameter bias and variability index for parameter bias, none of the interaction term had both the statistical significance and the practical significance.

Table D3.

Tests of between-subject effects on parameter bias computed for only filtered converged replications of the 2-class piecewise linear-linear LGMM.

		Statistical and Practical Significance	Manipulated Factors					
			Sample Size, n	Class Mixing Proportion, φ_k	Class Separation Location of Knot, γ_k	Mean of Slope of Second Phase, α_{4k}	Variance of Slope of Second Phase, $\sigma_{\eta_{4k}}^2$	Residual Variance σ_{ϵ}^2
Parameter Bias (Class 1)	Bias_ α_{11}	p-value =	0.000	0.010	0.000	0.000	0.000	0.000
		partial η^2 =	0.005	0.001	0.009	0.012	0.004	0.002
	Bias_ α_{21}	p-value =	0.000	0.044	0.000	0.000	0.002	0.035
		partial η^2 =	0.004	0.000	0.005	0.012	0.001	0.001
	Bias_ α_{41}	p-value =	0.173	0.000	0.003	0.179	0.796	0.363
		partial η^2 =	0.000	0.001	0.000	0.000	0.000	0.000
	Bias_ γ_1	p-value =	0.208	0.347	0.000	0.267	0.710	0.336
		partial η^2 =	0.000	0.000	0.002	0.000	0.000	0.000
	Bias_ $\sigma_{\eta_{11}}^2$	p-value =	0.564	0.002	0.000	0.021	0.000	0.028
		partial η^2 =	0.000	0.001	0.002	0.001	0.029	0.000
	Bias_ $\sigma_{\eta_{21}}^2$	p-value =	0.032	0.013	0.498	0.011	0.000	0.017
		partial η^2 =	0.001	0.001	0.000	0.001	0.040	0.000
Bias_ $\sigma_{\eta_{41}}^2$	p-value =	0.029	0.002	0.335	0.000	0.000	0.292	
	partial η^2 =	0.001	0.001	0.000	0.002	0.041	0.000	
Bias_ $\sigma_{\epsilon_1}^2$	p-value =	0.000	0.090	0.000	0.000	0.000	0.000	
	partial η^2 =	0.010	0.000	0.007	0.003	0.006	0.016	
Parameter Bias (Class 2)	Bias_ α_{12}	p-value =	0.000	0.024	0.000	0.000	0.000	0.000
		partial η^2 =	0.010	0.000	0.003	0.006	0.007	0.009
	Bias_ α_{22}	p-value =	0.002	0.696	0.113	0.000	0.000	0.005
		partial η^2 =	0.000	0.000	0.001	0.002	0.002	0.001
	Bias_ α_{42}	p-value =	0.002	0.527	0.014	0.000	0.000	0.014
		partial η^2 =	0.001	0.000	0.001	0.014	0.001	0.001
	Bias_ γ_2	p-value =	0.067	0.695	0.000	0.119	0.483	0.594
		partial η^2 =	0.000	0.000	0.002	0.000	0.000	0.000
	Bias_ $\sigma_{\eta_{12}}^2$	p-value =	0.024	0.002	0.000	0.091	0.000	0.033
		partial η^2 =	0.001	0.001	0.002	0.000	0.035	0.000
	Bias_ $\sigma_{\eta_{22}}^2$	p-value =	0.952	0.478	0.978	0.880	0.000	0.309
		partial η^2 =	0.000	0.000	0.000	0.000	0.007	0.000
Bias_ $\sigma_{\eta_{42}}^2$	p-value =	0.002	0.081	0.566	0.000	0.000	0.006	
	partial η^2 =	0.001	0.000	0.000	0.001	0.046	0.001	
Bias_ $\sigma_{\epsilon_2}^2$	p-value =	0.000	0.000	0.002	0.000	0.003	0.000	
	partial η^2 =	0.002	0.002	0.001	0.001	0.000	0.038	

Table D4.

Tests of between-subject effects on variability index for parameter bias for only filtered converged replications of the 2-class piecewise linear-linear LGMM.

		Statistical and Practical Significance	Manipulated Factors					
			Sample Size, n	Class Mixing Proportion, φ_k	Class Separation Location of Knot, γ_k	Mean of Slope of Second Phase, α_{4k}	Variance of Slope of Second Phase, $\sigma_{\eta_{4k}}^2$	Residual Variance σ_{ϵ}^2
Variability Index (Class 1)	Bias_ α_{11}	p-value =	0.258	0.607	0.008	0.026	0.055	0.378
		partial η^2 =	0.000	0.000	0.001	0.001	0.000	0.000
	Bias_ α_{21}	p-value =	0.923	0.285	0.180	0.382	0.907	0.932
		partial η^2 =	0.000	0.000	0.001	0.000	0.000	0.000
	Bias_ α_{41}	p-value =	0.012	0.003	0.535	0.000	0.417	0.023
		partial η^2 =	0.001	0.001	0.000	0.003	0.000	0.000
	Bias_ γ_1	p-value =	0.001	0.119	0.000	0.014	0.121	0.710
		partial η^2 =	0.001	0.000	0.002	0.001	0.000	0.000
	Bias_ $\sigma_{\eta_{11}}^2$	p-value =	0.221	0.569	0.028	0.594	0.288	0.906
		partial η^2 =	0.000	0.000	0.001	0.000	0.000	0.000
	Bias_ $\sigma_{\eta_{21}}^2$	p-value =	0.269	0.929	0.331	0.621	0.568	0.491
		partial η^2 =	0.000	0.000	0.000	0.000	0.000	0.000
Bias_ $\sigma_{\eta_{41}}^2$	p-value =	0.503	0.583	0.439	0.605	0.441	0.561	
	partial η^2 =	0.000	0.000	0.000	0.000	0.000	0.001	
Bias_ $\sigma_{\epsilon_1}^2$	p-value =	0.000	0.041	0.000	0.028	0.410	0.754	
	partial η^2 =	0.002	0.000	0.007	0.001	0.000	0.000	
Variability Index (Class 2)	Bias_ α_{12}	p-value =	0.063	0.732	0.002	0.159	0.387	0.094
		partial η^2 =	0.000	0.000	0.001	0.000	0.000	0.000
	Bias_ α_{22}	p-value =	0.923	0.285	0.180	0.382	0.907	0.932
		partial η^2 =	0.000	0.000	0.001	0.000	0.000	0.000
	Bias_ α_{42}	p-value =	0.205	0.013	0.000	0.000	0.861	0.024
		partial η^2 =	0.000	0.001	0.002	0.004	0.000	0.000
	Bias_ γ_2	p-value =	0.000	0.974	0.000	0.002	0.036	0.471
		partial η^2 =	0.002	0.000	0.002	0.001	0.000	0.000
	Bias_ $\sigma_{\eta_{12}}^2$	p-value =	0.189	0.370	0.220	0.998	0.319	0.451
		partial η^2 =	0.000	0.000	0.000	0.000	0.000	0.000
	Bias_ $\sigma_{\eta_{22}}^2$	p-value =	0.991	0.338	0.465	0.590	0.473	0.505
		partial η^2 =	0.000	0.000	0.000	0.000	0.000	0.000
Bias_ $\sigma_{\eta_{42}}^2$	p-value =	0.855	0.196	0.716	0.557	0.454	0.410	
	partial η^2 =	0.000	0.000	0.000	0.000	0.000	0.000	
Bias_ $\sigma_{\epsilon_2}^2$	p-value =	0.009	0.037	0.000	0.252	0.540	0.442	
	partial η^2 =	0.001	0.000	0.005	0.000	0.000	0.000	

NOTE that in the tests of between-subject effects on parameter bias and variability index for parameter bias computed for only filtered converged replications of the 2-class piecewise linear-linear LGMM, none of the interaction term had both the statistical significance and the practical significance.

REFERENCES

- Akaike, H. (1987). Factor analysis and AIC. *Psychometrika*, 52, 317–332.
- Bates, D. M., & Watts, D. G. (1988). *Nonlinear regression analysis and its applications*. New York: Wiley.
- Bauer, D. J. (2007). Observations on the use of growth mixture models in psychological research. *Multivariate Behavioral Research*, 42, 757-786.
- Bauer, D. J., & Curran, P. J. (2003). Distributional assumptions of growth mixture models: Implications for the overextraction of latent trajectory classes. *Psychological Methods*, 8, 338-363.
- Bauer, D. J., & Curran, P. J. (2004). The integration of continuous and discrete latent variable models: Potential problems and promising opportunities. *Psychological Methods*, 9, 3-29.
- Bollen, K. A., & Curran, P. J. (2006). *Latent growth models: A structural equation perspective*. New Jersey: Wiley.
- Celeux, G., & Soromenho, G. (1996). An entropy criterion for assessing the number of clusters in a mixture model. *Journal of Classification*, 13, 195–212.
- Chou, C.-P., Yang, D., Pentz, M. A., & Hser, Y.-I. (2004). Piecewise growth curve modeling approach for longitudinal prevention study. *Computational Statistics & Data Analysis*, 46, 213-225.
- Cudeck, R. (1996). Mixed-effects models in the study of individual differences with repeated measures data. *Multivariate Behavioral Research*, 31, 371-403.

- Cudeck, R., & Harring, J. R. (2007). The analysis of nonlinear patterns of change with random coefficient models. *Annual Review of Psychology*, 58, 615-637.
- Cudeck, R., & Harring, J. R. (2010). Developing a random coefficient model for nonlinear repeated measures data. In S.-M. Chow, E. Ferrer, & F. Hsieh (Eds.), *Statistical methods for modeling human dynamics: An interdisciplinary dialogue*. New York: Routledge.
- Cudeck, R., & Klebe, K. J. (2002). Multiphase mixed-effects models for repeated measures data. *Psychological Methods*, 7, 41-63.
- Curran, P. J. (2003). Have multilevel models been structural equation models all along? *Multivariate Behavioral Research*, 38 (4), 529-569.
- Davidian, M., & Giltinan, D. M. (2003). Nonlinear models for repeated measurements: An overview and update. *Journal of Agricultural, Biological, and Environmental Statistics*, 8, 387-419.
- Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood estimation from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B*, 39, 1-38.
- Duncan, T. E., Duncan, S. C., & Strycker, L. A. (2006). *An introduction to latent variable growth curve modeling: Concepts, issues, and applications* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum.
- Enders, C. K. (2006). Analyzing structural equation models with missing data. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course*. Greenwood, CT: Information Age Publishing, Inc.

- Figueiredo, M. A. T., & Jain, A. K. (2002). Unsupervised Learning of Finite Mixture Models. *IEEE Transactions Pattern Analysis and Machine Intelligence*, 24, 3, 381-396.
- Gagné, P. (2004). *Generalized confirmatory factor mixture models: A tool for assessing factorial invariance across unspecified populations*. Unpublished doctoral dissertation, University of Maryland, College Park.
- Gagné, P. (2006). Mean and covariance structure mixture models. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course*. Greenwood, CT: Information Age Publishing, Inc.
- Goldstein, H. (2003). *Multilevel statistical models* (3rd ed.). London: Edward Arnold.
- Grün, B., & Leisch, F. (2006). Fitting finite mixtures of linear regression models with varying & fixed effects in R. In A. Rizzi & M. Vichi (Eds.), *Compstat 2006 — Proceedings in Computational Statistics*. Heidelberg, Germany: Physica-Verlag.
- Hamilton, J. (2009). *An investigation of growth mixture models when data are collected with unequal selection probabilities: A Monte Carlo study*. Unpublished doctoral dissertation, University of Maryland, College Park.
- Hancock, G. R., & Lawrence, F. R. (2006). Using latent growth models to evaluate longitudinal change. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course*. Greenwood, CT: Information Age Publishing, Inc.
- Harring, J. R., Cudeck, R., & du Toit, S. H. C. (2006). Fitting partially nonlinear random coefficient models as SEMs. *Multivariate Behavioral Research*, 41, 579-596.

- Haughton, D. (1997). Packages for estimating finite mixtures: a review. *The American Statistician*, 51, 194-205.
- Hoogland, J. J., & Boomsma, A. (1998). Robustness studies in covariance structure modeling: An overview and a meta-analysis. *Sociological Methods and Research*, 26, 329-367.
- Jedidi, K., Jagpal, H., & DeSarbo W. S. (1997). Finite-mixture structural equation models for response-based segmentation and unobserved heterogeneity. *Marketing Science*, 16, 39-59.
- Laird, N. M., & Ware, J. H. (1982). Random-s models for longitudinal data. *Biometrics*, 38, 963-974.
- Leroux, B. G. (1992). Consistent estimation of a mixing distribution. *Annals of Statistics*, 20, 1350-1360.
- Lubke, G., & Muthén, B. (2007). Performance of factor mixture models as a function of model size, covariate effects, and class-specific parameters. *Structural Equation Modeling: A Multidisciplinary Journal*, 14, 26-47.
- Mann, H. M. (2009). *Testing for differentially functioning indicators using mixtures of confirmatory factor analysis models*. Unpublished doctoral dissertation, University of Maryland, College Park.
- Marsh, L. C., & Cormier, D. R. (2001). *Spline regression models*. Thousand Oaks, CA: Sage.
- Marsh, L. C., Maudgal, M., & Raman, J. (1990). Alternative methods of estimating piecewise linear and higher order regression models using SAS software. *SUGI*, 15, 523-527.

- McLachlan, G., & Peel, D. (2000). *Finite mixture models*. New York: Wiley.
- Meredith, W., & Tisak, J. (1990). Latent curve analysis. *Psychometrika*, 55, 107-122.
- Muthén, B. (2001). Latent variable mixture modeling. In G. A. Marcoulides & R. E. Schumacker (Eds.), *New developments and techniques in structural equation modeling* (pp. 1-33). Mahwah, NJ: Lawrence Erlbaum Associates.
- Muthén, B. (2002). Beyond SEM: General latent variable modeling. *Behaviormetrika*, 29, 81-117.
- Muthén, B. (2003). Statistical and substantive checking in growth mixture modeling. *Psychological Methods*, 8, 369-377.
- Muthén, B. (2004). Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data. In D. Kaplan (Ed.), *Handbook of quantitative methodology for the social sciences* (pp. 345-368). Newbury Park, CA: Sage.
- Muthén, B., & Muthén, L. (2000). Integrating person-centered and variable-centered analyses: Growth mixture modeling with latent trajectory classes. *Alcoholism: Clinical and Experimental Research*, 24, 882-891.
- Muthén, L. K., & Muthén, B. O. (1998-2010). Mplus User's Guide. Sixth Edition. Los Angeles, CA: Muthén & Muthén.
- Muthén, B., & Shedden, K. (1999). Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics*, 55, 463-469.
- Nylund, K. L., Asparouhov, T., & Muthén, B. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling: A Multidisciplinary Journal*, 14, 535-569.

- Paxton, P., Curran, P., Bollen, K., Kirby, J., & Chen, F. (2001). Monte Carlo experiments: Design and implementation. *Structural Equation Modeling: A Multidisciplinary Journal*, 8, 287-312.
- Pierce, C. A., Block, R. A., & Aguinis, H. (2004). Cautionary note on reporting eta-squared values from multifactor ANOVA designs. *Educational and Psychological Measurement*, 64, 916-924.
- Preacher, K. J., Wichman, A. L., MacCallum, R. C., & Briggs, N. E. (2008). *Latent growth curve modeling*. Thousand Oaks, CA: Sage.
- R Development Core Team (2009). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
- Roeder, K., & Wasserman, L. (1997). Practical Bayesian density estimation using mixtures of normals. *Journal of the American Statistical Association*, 92, 894-902.
- Sapp, M. (2006). *Basic psychological measurement, research designs, and statistics without math*. Springfield, IL: Charles C. Thomas Publisher.
- Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, 6, 461-464.
- Seber, G. A. F., & Wild, C. J. (1989). *Nonlinear regression*. New York: Wiley.
- Sokal, R. R., and F. J. Rohlf. (1995). *Biometry: The principles and practice of statistics in biological research* (3rd ed.). New York: W.H. Freeman.
- Soromenho, G. (1993). Comparing approaches for testing the number of components in a finite mixture model. *Computational Statistics*, 9, 65-78.

- Frühwirth-Schnatter, S. (2006). *Finite mixture and markov switching models*. New York: Springer.
- Tofighi, D., & Enders, C. K. (2008). Identifying the correct number of classes in growth mixture models. In G. Hancock & K. Samuelson (Eds.), *Advances in latent variable mixture models*. Charlotte, NC: Information Age Publishing, Inc.
- Tolvanen, A. (2008). *Latent growth mixture modeling: A simulation study*. Unpublished doctoral dissertation, University of Jyväskylä.
- Willet, J., & Sayer, A. (1994). Using covariance structure analysis to detect correlates and predictors of individual change over time. *Psychological Bulletin*, 116, 363-380.